THE UNIVERSITY
of EDINBURGH

THE UNIVERSITY *of* EDINBURGH

Synthetic biology enabled cellular and cell-free biosensors for environmental contaminants

Xinyi Wan（万心怡）

Thesis submitted for the degree of Doctor of Philosophy

University of Edinburgh

School of Biological Sciences

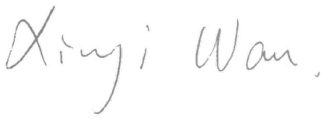
Year of Submission 2019

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification.

Xinyi Wan

06/03/2019



Acknowledgement

I would not be able to finish this work without the constant help and support I have received. A few words would never be enough to express my gratitude to the people who have been supportive. I would like to wish everyone the best luck for their life and career.

Great thanks to my supervisor Dr Baojun Wang for his guidance, supervision and encouragement. Without his support, I would not be able to start and finish my study in Edinburgh, let alone to explore the wonderful world in Synthetic Biology.

I would also like to thank my supervisor Prof Christopher French for his advice and patience, especially for his guidance during my difficult time and for his advice on my thesis.

Thank you also to my committee member Dr Ramon Grima for his advice and patience. Without his support, I would be much more stressed when I am writing these lines.

Many thanks to my examiners Dr Alistair McCormick and Prof Belkin Shimshon for spending their precious time to evaluate this work.

Thanks to Dr Morsaline Billah (Khulna University) and his colleagues for facilitating my collections of groundwater samples in Bangladesh.

Also, thank you to my colleagues and friends I have met in Edinburgh. Sally, Duan, Fei, Wei, Yang, Trevor, Russell, Filipe, Louise, Ella, Prabu, Yiyu, Grant, Jan, Marcos, Chaokuo, Julia, Matt, Smitha, Eric, Maryia, Maia, Annegret … It has been very delightful to have you all around and being supportive.

Special thank you to my husband and also my previous colleague Elvis for always being supportive, patient, understanding and up-lifting, especially for when I was stressed and in bad moods. Thank you also for your immediate advice on this work. Sorry that I always need to work overtime, but thank you for always taking care of me.

I would also like to thank my family. My father Fujin and my mother Kongyu have been very supportive at all times both mentally and financially for my study in Edinburgh. Also thank you to my grandfather Henghua, grandmother Fugying and all my other family members who have been generous and supportive.

Lastly, I would like to thank the sponsors of my PhD scholarship – School of Biological Sciences, The University of Edinburgh, China Scholarship Council and Scottish Universities Life Sciences Alliance.

Table of Contents

Lay Summary

Cell-based biosensors have great potential to detect various toxic and pathogenic contaminants in aqueous environments. However, frequently they cannot meet practical requirements due to insufficient sensing performance, inadequate sensing platforms and biosafety issues. My PhD project aims to address these challenges to allow cell-based biosensors to achieve their projected applications in the field.

I first investigated a general methodology which can significantly improve the sensing performance for a variety of cell-based biosensors. This methodology combines multiple synthetic biology strategies, which can optimise sensor sensitivity, maximum output levels and background levels step by step. I applied this methodology to previously developed bacterial sensors for sensing arsenic and mercury, and I successfully improved their sensitivity and outputs up to 5,000-fold and 750-fold respectively.

Facilitated by this methodology, I created a sensor library, with each sensor exhibiting a different sensitivity to the target. Using these sensors, I developed easy-to-use sensing platforms for cost-effective and portable field testing, where the target signals were easily quantified by simple visualisation. Physical entrapment methods were applied to the sensing platforms to mitigate and minimise the biosafety concerns.

Finally, to further minimise the biosafety issues, the arsenic and mercury sensors were transferred into a cell-free system which does not contain living cells. Aforementioned methodology combined with additional tuning methods were applied to improve sensors in this system. A paper-based cell-free sensing platform was also developed and tested. A mercury sensor with colorimetric output was adapted to this system, where the sensor's response to 2 ppb mercury could be easily seen by the naked eye.

Overall, the ultrasensitive signal amplifying methodology along with the cellular and CFS-based sensing platforms can be widely applicable to many other cell-based sensors, thus paving the way for their projected applications in the real world.

Abstract

Cell-based biosensors have great potential to detect various toxic and pathogenic contaminants in aqueous environments. However, frequently they cannot meet practical requirements due to insufficient sensing performance, inadequate sensing platforms and biosafety issues.

Here, I investigated a novel, comprehensive and modular methodology for optimising cell-based biosensors to address these challenges, and to enable them for their practical applications. In particular, this methodology combines multiple synthetic biology strategies, which can systematically and significantly improve sensors' sensing performance in a predictable manner. It first optimises a sensor's sensitivity by regulating its intracellular receptor densities, then further improves its output by applying a multi-layer transcriptional amplifier cascade, and finally regulates its leakiness by combining promoter structure engineering and post-translational regulation. Exemplary bacterial cell-based arsenic and mercury sensors were used to demonstrate this methodology, and their detection limits and outputs were improved up to 5,000-fold and 750-fold respectively.

Facilitated by this methodology, I developed easy-to-interpret sensing platforms for cost-effective and portable field testing, where the analytes were easily quantified by simple visualisation. Physical entrapment methods, i.e., agarose gel entrapment and microfluidic biodisplay, were applied to the sensing platforms to mitigate and minimise the biosafety concerns.

To further eliminate the biosafety issues, the arsenic and mercury sensors were transferred into a crude cell extract-based cell-free system (CFS). To adapt the sensors to the CFS, aforementioned methodology combined with additional tuning methods were applied, such as tuning the sensors' DNA concentration and their receptor to promoter ratio, introducing transcriptional amplifiers and promoter engineering. A similar paper-based sensing platform could be generated based on these optimisation methods. A mercury sensor with colorimetric output was adapted to a paper-based CFS, where the sensor's response to 2 ppb mercury could be easily visualised by the naked eye.

Overall, the verified signal amplifying methodology along with the cellular and CFS-based sensing platforms can be widely applicable to many other cell-based sensors, paving the way for their real world applications in the environment and healthcare.

Abbreviations

AAS	Atomic absorption spectroscopy
ABS	ArsR binding site
AHL	[N-]acyl-homoserine lactones
aTc	Anhydrotetracycline
ATP	Adenosine-5'-triphosphate
CAMERA	CRISPR-mediated analog multi-event recording apparatus
Cas	CRISPR associated system
CE-CFS	Cell extract-based cell-free system
CFS	Cell-free system
CPRG	Chlorophenol red-β-D-galactopyranoside
CRISPR	Clustered regularly interspaced short palindromic repeats
DMF	N,N-dimethylformamide
DNA	Deoxyribonucleic acid
EPA	Environmental protection agency
FACS	Fluorescence-activated cell sorting
Ficoll	Copolymer of sucrose and epichlorohydrin
GFP	Green fluorescent protein
GSH	Glutathione
hrp	Hypersensitive response and pathogenicity
ICP-MS	Inductively-coupled plasma mass spectrometry
IMBED	Ingestible micro-bio-electronic device
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Luria-Bertani broth
LOD	Limit of detection
NO	Nitric oxide

PCR	Polymerase chain reaction
PEG	Polyethylene glycols
PURE	Purified recombinant elements
ppb	Parts per billion based on weight/volume (μg/L)
R	Receptor
RBS	Ribosome binding site
RFP	Red fluorescent protein
RNA	Ribonucleic acid
RNAP	RNA polymerase
σ^{54}	RNAP sigma-54 factor (RpoN)
σ^{70}	RNAP sigma-70 factor (RpoD)
TA	Transcriptional activator
TA-Amp	Activator-based amplifier
TAmp	Transcriptional amplifier
TF	Transcription factor
TR	Transcriptional repressor
TR-Amp	Repressor-based amplifier
TX-TL	Transcription-translation
WHO	World Health Organization
X-gal	5-bromo-4-chloro-3-indolyl β-D-galactopyranoside
YFP	Yellow fluorescent protein

List of Figures

List of Tables

Chapter 1. Introduction

Whole-cell biosensors are cells that can detect and report a target or condition of interest (Daunert et al. 2000; van der Meer and Belkin 2010; Wang and Buck 2012; Gui et al. 2017). They can be either domesticated cells that naturally have sensing and reporting functions, or genetically modified cells that have been endowed with desired functions. Compared with conventional physical and chemical detection methods, whole-cell biosensors are cost-effective (e.g., reproducible and easy to grow), simple to manufacture (e.g., no need for sophisticated equipment) and environmentally friendly (e.g., no toxic chemical reactions and biodegradable). Both prokaryotic and eukaryotic biosensors have been developed; the former is a particularly attractive option as they are generally amenable to genetic manipulation and systems design (van der Meer 2010). Due to these advantages, a number of prokaryotic biosensors have been studied for various purposes, such as environmental assessment (Stocker et al. 2003; De Mora et al. 2011; Wang et al. 2013; Huang et al. 2015b; Hwang et al. 2016; Kim et al. 2016; Cao et al. 2017; Cayron et al. 2017), clinical diagnosis (Duan and March 2010; Saeidi et al. 2011; Archer et al. 2012; Gupta et al. 2013; Kotula et al. 2014; Hwang et al. 2014; Courbet et al. 2015; Danino et al. 2015; Cayron et al. 2017; Daeffler et al. 2017; Riglar et al. 2017; Ho et al. 2018; Watstein and Styczynski 2018; Mimee et al. 2018), controlled bioprocessing (Zhang and Keasling 2011; Zhang et al. 2012), mineral surveying (Cerminati et al. 2011) and landmine clearing (Belkin et al. 2017).

Despite the advantages and demonstrated successes in the laboratory, very few prokaryotic biosensors have been successfully commercialised. Biosafety concerns are the most important factor that restricts the employment of cellular sensors (Dana et al. 2012). Nevertheless, poor sensing performance in the natural environment may also pose problems: **1)** Low selectivity would make sensors less accurate in a complicated environment (Kim et al. 2016; Shemer et al. 2017; Landry et al. 2018). **2)** Inadequate output would make them less effective in reporting the presence of targets (Wang et al. 2014; Wang et al. 2015). **3)** High background leakage may reduce the robustness of the reporting output (e.g., can be easily saturated and exhibit a lower dynamic range) (Wackwitz et al. 2008; Fernandez-Rodriguez and Voigt 2016; Merulla and van der Meer 2016). Finally**, 4)** the lack of cheap and easy-to-use sensing platforms for cell-based biosensors prevents them being accepted by the general population (Prindle et al. 2012; Volpetti et al. 2017). Fortunately, recent developments in synthetic biology have provided numerous solutions.

Synthetic biology is an emerging field that applies established principles from engineering to biology with the aim to achieve rational design of biological systems (Endy 2005; Purnick and Weiss 2009; Cheng and Lu 2012; Wang and Buck 2012; Way et al. 2014; Bradley et al. 2016a; Bradley et al. 2016b; Bashor and Collins 2018). It uses catalogued and standardised biological components to redesign natural biological systems, alongside design and construction of new biological parts, devices, and systems with predictable, useful and novel functions. With over 50 years' research on molecular biology and functional genomics, an inventory of biological parts has been compiled, allowing for rational design in synthetic biology (Weber and Fussenegger 2012; Cameron et al. 2014). Moreover, with recent advances in technology, it is easy to analyse, synthesise, assemble, modify and transfer genetic components into living organisms (Weber and Fussenegger 2012; Church et al. 2014; Cameron et al. 2014).

For my doctoral study, I focused on using synthetic biology tools and strategies to optimise cellular biosensors in order to address the aforementioned issues to meet their applications requirements. I tested and validated those strategies on previously studied bacterial sensors for sensing arsenic and mercury. Arsenic and mercury contamination in drinking water are environmental and health issues that continue to threaten the life of millions of people globally. I aimed to improve the sensors to meet their application criteria and also to develop inexpensive, portable and easy-to-use sensing platforms to facilitate their application in the field.

In the following sections, I will introduce the background of my research by describing : **1)** the current situation of arsenic and mercury contamination in the environment and their traditional detection methods (Section 1.1); **2)** the history of biosensors and principles of designing prokaryotic cell-based biosensors (Section 1.2); **3)** how to use synthetic biology tools to improve the performance of prokaryotic cellular biosensors (Section 1.3); **4)** how to expand their functions for specialised application requirements (Section 1.4); **5)** current issues and possible solutions regarding the use of biosensors in the field (Section 1.5).

1.1 Arsenic and mercury contamination

1.1.1 Arsenic occurrence and toxicity

Arsenic (As) is a metalloid element with atomic number 33 and atomic weight 74.92 g/mol. It exists naturally in the Earth's crust, and is widely distributed throughout the air, water and land. Normally, arsenic concentration in water is less than 1 – 2 μg/L. However, due to sulfide mineral deposits and sedimentary deposits derived from volcanic rocks, arsenic concentration can be significantly elevated, particularly in ground water (Bhattacharjee and Rosen 2007; WHO 2011). Additionally, arsenic can be introduced into water by industrial effluents such as

mining wastes (Hindmarsh and McCurdy 1986). The arsenic exists in oxidation states of -3, 0, +3 and +5, and it mostly forms arsenic sulphide, metal arsenates or arsenides (WHO 2011). In oxygenated surface water, arsenate (e.g., AsO_4^{3-}) is generally the most common arsenic species, while under reducing condition such as deep groundwater, the predominant arsenic species is arsenite (e.g., AsO_2^-) (Sharma and Sohn 2009; WHO 2011). Reduction to arsine (AsH_3) and elemental arsenic is very rare in nature (Sharma and Sohn 2009).

In the last few decades, arsenic has drawn attention worldwide for its toxicity to humans. The acute toxicity of arsenic compounds to humans is predominantly a function of their rate of removal from the body. Arsine is considered to be the most toxic form of arsenic, followed by the arsenites (As^{3+}), arsenates (As^{5+}) and inorganic arsenic (WHO 2011). Arsenite can bind to intracellular sulphur thiolates and glutathione (GSH), leading to the depletion of the cellular GSH pool and the disruption of cellular redox status (Rodríguez et al. 2005; Sattar et al. 2016). Moreover, arsenite can inhibit the pyruvate oxidation pathway and tricarboxylic acid cycle (Schiller et al. 1977). It has also been proved that arsenite can indirectly exacerbate deoxyribonucleic acid (DNA) damage (Qin et al. 2008). The toxicity of arsenate is due to its similarity with phosphate, which disturbs phosphate pathways and leads to reduced production of adenosine-5'-triphosphate (ATP) (Sattar et al. 2016).

Acute arsenic poisoning via ingestion of highly contaminated water is associated initially with nausea, vomiting, abdominal pain and severe diarrhoea (Ratnaike 2003; WHO 2011). Chronic arsenic intoxication may cause skin lesions (e.g., hyperpigmentation and hypopigmentation), peripheral neuropathy, vascular diseases, and cancers in skins, brain, kidney, lung, stomach and bladder (Sharma and Sohn 2009; WHO 2011).

Due to the high risk of long-term exposure to arsenic, the World Health Organisation (WHO) has set a provisional guideline value of 10 μg/L (10 ppb) of arsenic in drinking water. It has been reported by WHO that more than 50 countries with more than 140 million people have been drinking water with higher than 10 μg/L arsenic contamination (Ravenscroft et al. 2009). In such highly arsenic contaminated areas, the effect can be devastating. For example, in Bangladesh it was estimated there were nearly 43,000 deaths related to arsenic poisoning per year (Flanagan et al. 2012).

1.1.2 Mercury occurrence and toxicity

Mercury (Hg) is a metallic element with atomic number 80 and atomic weight 200.59 g/mol. It occurs naturally in the environment, and is widely distributed through natural processes such as volcanic activity and weathering of rocks (WHO 2011). However, the main release of mercury into the environment is due to human activities, most significantly after the industrial

revolution in the 19th century (WHO 2011). Such activities include coal-fired power stations, residential coal burning, waste incinerators, industrial processes of mining for mercury and other metals, and production of electrolytic and electrical appliances; also, various mercury compounds have been used as fungicides, preservatives, pharmaceuticals, electrodes, dental amalgams and some ethnic and folk remedies (WHO 2011). Normally the inorganic mercury level in surface water and groundwater is below 0.5 μg/L (WHO 2011). However, mercury-related human activities and frequent volcanic activity can significantly increase the mercury level in groundwater. It has been reported that among 6,856 samples of groundwater from the state of California, 225 samples contained 0.21-300 μg/L of mercury, and 27 samples of them exceeded 2 μg/L (2 ppb, a guideline set by the US Environmental Protection Agency (EPA) for drinking-water) (Park and Zheng 2012). Multiple rivers in China have also been reported to contain higher than 5 μg/L of mercury (Zhang and Wong 2007). Up to 5.5 μg/L of mercury in wells has been reported in Japan where volcanic activity is frequent (WHO 2005).

It has been widely reported that mercury poisoning can affect the human central and peripheral nervous, renal, immune, cardiovascular systems (WHO 2005; Bernhoft 2012). Mercury in all forms can change the structure and inactivate proteins; it also binds to sulfhydryl group leading to membrane degradation and cell death (Dhuldhaj et al. 2012). These will result in cellular dysfunction and tissue damage. Acute exposure to inorganic mercury (e.g., $HgCl_2$) can cause haemorrhagic gastritis and colitis, acute renal failure and death (WHO 2005; Bernhoft 2012). Absorption of organic mercury (e.g., methyl mercury and ethyl mercury from the food chain) is reported to be the main cause of neurological damage, leading to Alzheimer's and Parkinson's diseases in adults and learning disabilities in children (Olivieri et al. 2002; WHO 2005; Zahir et al. 2005). To protect humans from being exposed to mercury, several methods have been used to reduce the use and release of mercury in industrial processing. For inorganic mercury, WHO has set a guideline of 6 μg/L for drinking water.

1.1.3 Traditional detection methods and their drawbacks

To maintain an acceptable quality of drinking water for human health, arsenic and mercury levels need to be monitored regularly. Detection methods need to be accurate and sensitive enough to meet the guidelines. Traditional methods are based on the physical properties of target elements, such as atomic absorption spectroscopy (AAS) and inductively-coupled plasma mass spectrometry (ICP-MS). The detection limit of these methods ranges from 0.1 μg/L to 2 μg/L for arsenic and from 0.05 μg/L to 5 μg/L for mercury (WHO 2011). However, these techniques can be difficult in terms of sample preparation, equipment operation and data analysis due to the specialist training required for each step (WHO 2011; Wei et al. 2014; Kaur

et al. 2015). Moreover, the instrument and reagents themselves are expensive. For example, £ 50 per hour is normally charged to use ICP-MS for element analysis (approximately £ 5 per sample). For AAS, the price is about £ 5 per sample per element. Most importantly, the sophisticated equipment requires lab-based use, making these methods unsuitable for on-site testing.

Chemical test kits have also been developed for water quality testing. They do not require expensive instruments, and some kits are very sensitive and time efficient, particularly for arsenic (the current kits for mercury are only for highly contaminated samples; for example, the Osumex HMT Mercury Test Kit can only detect ≥ 25 μg/L of mercury). As an example, the well-known Gutzeit method is fast (about 5 min), sensitive (detection limit ≤ 2 μg/L), easy to operate, and the arsenic concentration is determined by colorimetric output which can be observed by the naked eye (Sanger 1908; Salman et al. 2012). However, the chemical reactions for heavy metals always involve toxic reagents or products. For example, the Gutzeti reaction requires concentrated hydrochloric acid which is highly corrosive; it also needs mercuric bromide solution and produces arsine gas which are very toxic. Therefore, the chemical methods may cause issues for transport and handling (Das and Sarkar 2016). Based on the Gutzeti method, Merck (Germany) and Hach (USA) have developed arsenic test kits (Merckoquant® Arsenic, product No. 1.17927.0001, and EZ arsenic test kit, product No. 2822800) which are commercially available for detecting arsenic in groundwater. Both kits use solid sulfamic acid instead of concentrated hydrochloric acid, but this reduces the sensitivity to 5 – 10 μg/L, and the response time is increased to 20 – 35 min (Das and Sarkar 2016). Another disadvantage of these kits is that they are still based on visualising colour scales, which can make it difficult to determine an accurate quantitative result. Kearns and Tyson have digitised the Hach EZ arsenic kit by analysing the colour developed on the detector strip (Kearns and Tyson 2012). The digitised results can be more accurate, but the results processing required 24 h which makes it more time consuming. The Arsenator® Digital Arsenic Test Kit developed by Wagtech (UK) is also a well-known arsenic detection kit, and can detect 2 – 100 μg/L of arsenic in 20 min. However, the cost of this kit ($ 4.95 per test) is 2 – 5 times higher than the other kits ($ 1 – 2 per test) (Das and Sarkar 2016). Recently, an enzymatic reaction-based arsenic testing kit (http://www.bbsrc.ac.uk/news/health/2018/180222-n-simple-arsenic-sensor-could-save-lives/) was developed by researchers from University College London and Imperial College London. It is cheap and easy to use, and can produce a digital reading in one minute upon sensing arsenic; however, the stability and shelf-life of the enzymes may need to be evaluated.

1.2 Biosensors

Biosensors represent an alternative method for detecting threats for human beings. The earliest warning system can be traced back to the late 1890s, when it was first proposed by a Scottish physiologist to use canaries to detect carbon monoxide in coal mines (Wikipedia 2018a; Wikipedia 2018b). Canaries are particularly sensitive to toxic gases such as carbon monoxide, which is colourless, tasteless and odourless and often occurs underground during a mine fire or after an explosion. Any sign of distress from canaries is a clear warning signal to show that the underground environment is unsafe, and miners should be evacuated from this place for their own safety (BBC News 1986). In addition to canaries, different animals have been used as a detector for different purpose; for example, 'dancing cat fever' was a sign of eating mercury-contaminated fish (Kurland et al. 1960), and pigeons have been used as a biological indicator for urban lead pollution (Ohi et al. 1981). All of these examples reflect the fact that living organisms have their own sensing systems for particular environmental signals, and some are more sensitive than the others. However, toxicological and eco-toxicological testing on animals is a complicated procedure and is usually time consuming and involves ethical issues. To replace the animal models, cell lines and microorganisms have become popular.

Microorganisms would be a particularly attractive option because they are naturally exposed to various environments and have developed their own systems to sense and respond to specific environmental signals. Although those sensing systems can be either simple transcriptional networks or more complicated signalling pathways, they share similar architecture: recognising external signals via sensing modules, transducing them into intracellular signals via regulatory systems, and executing physiological responses through output actuators (**Fig. 1.1a**) (van der Meer and Belkin 2010; Wang and Buck 2012; Kim et al. 2018). More importantly, microorganisms are generally amenable to genetic manipulation and systems design. With the development of synthetic biology, it is becoming more achievable to design more complex systems in microorganisms to perform desired functions.

1.2.1 Natural prokaryotic biosensors and sensing modules

Before the advances in genetic engineering, the development of biosensors mainly relied on serendipity. Prokaryotes have natural abilities to sense environmental signals, such as chemicals, proteins, pathogens, light, temperature and radiation etc. Moreover, they can respond to these signals by performing specific actions, such as changes in oxygen respiration rate, cytoplasmic enzyme activities, motility, growth and mutagenicity (van der Meer 2010).

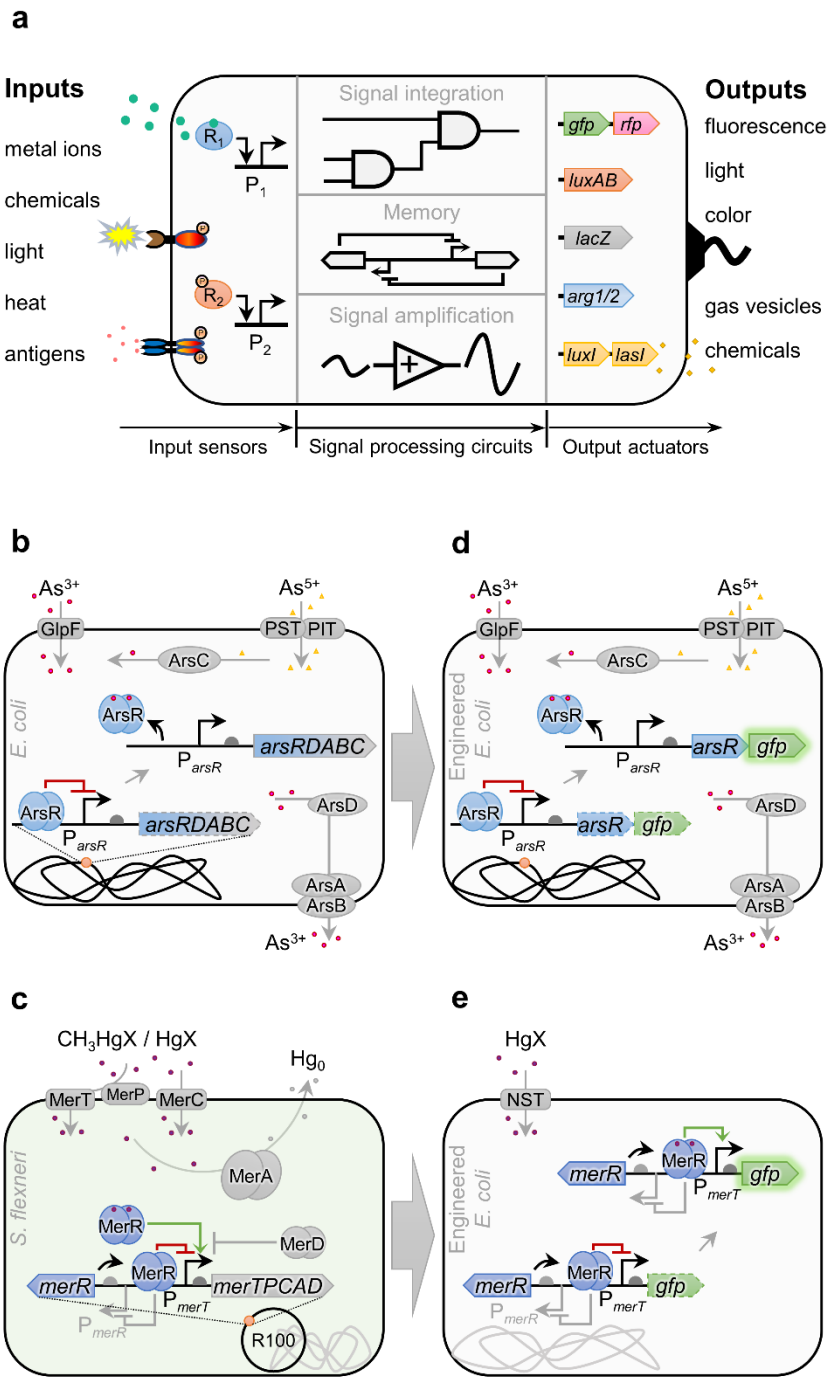


Figure 1.1: Architecture and engineering of synthetic biosensors.

(**a**) Architecture of a modular synthetic biosensor. R, receptor. P, promoter. *gfp*, gene encoding a green fluorescent protein. *rfp*, gene encoding a red fluorescent protein. *luxAB*, genes encoding a bacterial luciferase for luminescent output. *lacZ*, gene coding β-Galactosidase for colorimetric output. *arg*, acoustic reporter genes which express gas vehicles that are detectable by ultrasound. *luxI* & *lasI*, genes encoding synthases for quorum sensing molecules. (**b**) *ars* operon from *E. coli*'s chromosome and its role in arsenic regulation (Silver and Phung 2005; Chen and Rosen 2014). GlpF, an aquaglyceroporin. PST, phosphate-specific transport system. PIT, phosphate inorganic transport system. (**c**) *mer* operon

One well-known test is the commercialised Microtox™-test developed by Bulich and Isenberg (Bulich and Isenberg 1981), in which the marine bioluminescent bacteria *Allovibrio fischeri* were tested with different samples and chemicals while their bioluminescence activity was constantly measured (van der Meer 2010). Because bioluminescence is a highly energy demanding process in the cells, any compounds that disrupt this process will reduce luciferase activity which leads to the decrease of light output in this assay. Based on this finding, many other studies have used bioluminescent bacteria as reporters for toxic environment. However, this natural light change is not particularly responsive to any one specific compound, which eliminates the possibility to use it for sensing and reporting a toxic signal selectively in a complex environment. Therefore, specific sensing systems will be more ideal for sensing and reporting a unique target.

One popular sensing system for testing toxic environments is the heavy metal sensing system. As many heavy metals, such as arsenic and mercury, naturally exist in the environment, it is very likely that some bacteria are living in an environment with those heavy metals. Because the heavy metals are toxic to living organisms, including bacteria, they have developed specific systems to sense and detoxify those heavy metals. Under extreme conditions, although controversial, bacteria may even use heavy metals to process essential cellular functions (Wolfe-Simon et al. 2011). Thus, many heavy metal-responsive bacteria have been isolated from highly contaminated environments and sewage.

As arsenic is widely distributed in the environment, many bacteria have been found to be resistant to arsenic and some can also 'eat and breath' arsenic (i.e., use arsenic as electron donor or acceptor for respiration) (Cervantes et al. 1994; Nealson et al. 2002; Silver and Phung 2005; Andres and Bertin 2016). The resistance against arsenic is provided by the *ars* operon from the bacterial chromosome, plasmid or both (Andres and Bertin 2016). A well-defined *ars* system from *Escherichia coli* is briefly described in **Figure 1.1b**. In both gram-negative and gram-positive bacteria, ArsR, the arsenic (particularly As^{3+}) receptor is the first protein in the resistance system to recognise and bind to the arsenic. ArsR is a transcriptional repressor, which represses the P_{arsR} promoter by binding it with its C-terminus and preventing the access of ribonucleic acid (RNA) polymerase (RNAP) (Shi et al. 1996; Saha et al. 2017). The whole *ars* operon is under the control of this promoter. Arsenic binding changes the conformation of

from *Shigella flexneri* R100 plasmid and its role in mercury regulation (Misra et al. 1985; Barkay et al. 2003). CH_3HgX, organic form of mercury. HgX, inorganic form of mercury. (**d**) An engineered *E. coli* biosensor with arsenic sensing and reporting function. (**e**) An engineered *E. coli* biosensor with mercury sensing and reporting function.

ArsR, releases the ArsR from the P_{arsR} promoter, and as a consequence activates the expression of the *ars* operon (e.g., *arsRDABC* in *E. coli*) and hence, by negative feedback, upregulates the *arsR* expression. The ArsR and P_{arsR} promoter, therefore, form the sensing module that controls arsenic resistance in bacteria. Although ArsR can also be regulated by antimonite and bismuth, it has been mainly used for sensing arsenic as the others are usually at very low levels in the environment (Stocker et al. 2003; Trang et al. 2005; Joshi et al. 2009; De Mora et al. 2011; Siegfried et al. 2012; Wang et al. 2013; Huang et al. 2015a; Li et al. 2015). Like ArsR, other ArsR-SmtB family proteins (e.g., CadC, AztR, ZiaR, CmtR, CzrA, NmtR, KmtR and BxmR) and their cognate promoters can respond to cadmium, lead, zinc, cobalt, nickel, copper and silver in a similar way in bacteria (Shi et al. 1996; Saha et al. 2017).

Like arsenic resistance, bacteria have also evolved resistance systems to mercury. The bacterial resistance to mercury was first found in a clinical isolate of *Staphylococcus aureus* by Moore *et al.*, and it is widely transferred among microbial communities via plasmids (Moore 1960; Barkay et al. 2003; Mathema et al. 2011). The genetic system of mercury resistance is present as the *mer* operon, which is known to be the only bacterial metal resistance system with large-scale transformation of its toxic target to non-toxic forms (Barkay et al. 2003; Mathema et al. 2011). Therefore, there has been a long-time interest in the role of the *mer* operon in global recycling of mercury, and utilising it for mercury bioremediation. These also motivated the mercury to be the first specific target of the genetically engineered whole-cell biosensors for heavy metals (Selifonova et al. 1993). Although the first bacteria found with mercury resistance were Gram-positive, the *mer* operon has been studied more extensively in Gram-negative bacteria (Mathema et al. 2011). The function of the *mer* operon of Gram-negative bacteria is summarised in **Figure 1.1c**. Although mercury can enter the cell without any transport system, bacteria nevertheless evolved a mercury uptake system to reduce its toxicity to other cysteine-rich proteins (Barkay et al. 2003). Organic and inorganic mercury can be transported into cells by a small periplasmic protein MerP and inner membrane proteins like MerT and MerC. Enzymes MerA and MerB are involved in the volatilisation and release of Hg^0 from inorganic and organic mercury respectively. MerR is a transcriptional regulator in the *mer* operon to control its activity. Unlike ArsR-SmtB family regulators, MerR family proteins (e.g., MerR, ZntR, PbrR, CueR, GolS and CoaR) are defined as transcriptional activators in bacteria for sensing relevant metals (except MerR from *Streptomyces lividans* 1326, which reportedly functions as a repressor like ArsR-SmtB family regulators) (Shi et al. 1996; Saha et al. 2017). When there is no Hg^{2+} in the cells, MerR will bind to its cognate promoters P_{merR} and P_{merT} and repress them. When there is a high concentration of Hg^{2+}, MerR will bind to Hg^{2+} and aid the activation of the promoters by a conformation change.

Here, MerR and its cognate promoters can be considered as the sensing system of mercury. Most importantly, MerR is very specific to Hg^{2+} *in vivo*, so it is an ideal candidate for building mercury biosensors (Virta et al. 1995; Barkay et al. 2003; Pellinen et al. 2004; Wang et al. 2013; Wang and Buck 2014; Didovyk et al. 2017).

Apart from sensing heavy metals, prokaryotes have developed other systems to respond to other environmental signals. For example, the LacI system (targeting isopropyl β-D-1-thiogalactopyranoside (IPTG) and lactose), AraC system (targeting L-arabinose) and TetR system (targeting tetracycline) from *E. coli* have been widely used for demonstrating genetic circuits for synthetic biology (van der Meer and Belkin 2010; Mahr and Frunzke 2016). Prokaryotes also have developed sensing systems for detecting toxic compounds or warning chemicals; for example, the *IbpR* system from *Pseudomonas putida* can be used to detect various aromatic pollutants in the environment (Selifonova and Eaton 1996), and the YhaJ-P_{yqjF} system from *E. coli* has been used to report DNT/TNT from landmines (Belkin et al. 2017). Light-responsive switches are also becoming popular for controlling synthetic cellular pathways due to their faster response compared to traditional chemical inducers (Olson et al. 2014). The CcaS-CcaR system and Cph8-OmpR systems from *Synechocystis* sp. and YF1-FixJ system from *Bradyrhizobium* have been used for sensing green, red and blue light respectively in *E. coli* (Olson et al. 2014; Fernandez-Rodriguez et al. 2017).

Pathologic signal sensing has also been generated from bacteria. These can be based on quorum sensing systems, such as LasR system (targeting 3-oxo-C12-HSL, a type of [N-]acyl-homoserine lactone (AHL)) from *Pseudomonas aeruginosa* can be used to sense this pathogen and has been engineered to report and kill this pathogen (Saeidi et al. 2011; Gupta et al. 2013; Hwang et al. 2014). Nitric oxide (NO, an inflammation biomarker) responsive systems such as the NorR-P_{norV} and NsrR-P_{yeaR} evolved for NO detoxification from *E. coli* have been used for inflammation diagnosis in the human gut, urine and blood (Lin et al. 2007; Archer et al. 2012; Courbet et al. 2015). Other inflammation biomarkers, such as colonic thiosulfate and tetrathionate, can be detected by the ThsSR and the TtrSR systems derived from *Shewanella sp.* (Daeffler et al. 2017) or *Salmonella typhimurium* (Riglar et al. 2017). Bacteria, such as *Salmonella*, *Escherichia*, *Clostridium* and *Bifidobacterium*, have also been used to target tumours due to their natural ability to accumulate in hypoxic environments (as in most tumours) and invade tumour cells (Forbes 2010; Weber and Fussenegger 2012; Danino et al. 2015). Moreover, hypoxia responsive systems such as the P_{fdhF} promoter derived from *E. coli* (Anderson et al. 2006) and the P_{pepT} promoter from *Salmonella* SL7207 (Yu et al. 2012) have been engineered in other bacteria to enhance their accumulation at tumour sites.

In conclusion, bacteria have been found to naturally respond to various signals from the environment. These natural sensing systems can be specific to a unique target, and can also be active in the presence of a group of similar targets. In some circumstances, the microorganisms naturally harbouring such sensing systems may be directly used as sensors. However, some microorganisms are difficult to work with, and most response behaviour naturally connected to the sensing systems cannot be directly monitored, or are not useful for application purposes. Currently, with the rise of synthetic biology, people tend to extract the central parts (e.g., promoters and regulators) from these sensing systems, use them as sensing modules and connect them with desired response behaviour in an easy-to-use bacterium.

1.2.2 Diverse output modules

The simplest prokaryotic biosensor can be easily generated by directly connecting a sensing module with an output module. For example, the first reports of genetically engineered bacteria were generated by fusing an colorimetric output to a bacterial native operon which was responsible for mutagenesis (Quillardet et al. 1982; Oda et al. 1985). Various proteins have been used in output modules for biosensors, such as the proteins that produce light, fluorescence, colours, electrons or gas vesicles for monitoring, or special regulatory proteins related to cell movement and killing effects (**Fig. 1.1a**). The output proteins should be chosen carefully based on sensor applications. Specially, for reporting and monitoring a target, the output proteins should be selected based on their sensitivities and the convenience of measurement (Kim et al. 2018).

1.2.2.1 Luciferase

Bioluminescent output is based on enzymatic reaction which produces light source without the need of excitation. It has been used in the first bacterial whole-cell biosensor for detecting a specific chemical (naphthalene) (King et al. 1990), and it has been continuously applied to other sensors development due to their low background and high sensitivity (Virta et al. 1995; Reid et al. 1998; Stocker et al. 2003; Pellinen et al. 2004; Michelini et al. 2008; Roda et al. 2011; Siegfried et al. 2012; Danino et al. 2015; Huang et al. 2015b; Cevenini et al. 2018; Mimee et al. 2018). Bacterial (LuxCDABE or LuxAB) (Reid et al. 1998; Stocker et al. 2003; Danino et al. 2015; Huang et al. 2015b; Mimee et al. 2018) and firefly (LucFF) (Virta et al. 1995; Pellinen et al. 2004; Roda et al. 2011) luciferases are often used as sensor outputs. Luciferase from other organisms (e.g., click beetle, copepod and sea shrimp) and luciferase mutants with different emission spectra and improved activity have been discovered and developed (Loening et al. 2007; Roda et al. 2009; Mary P Hall et al. 2012; Roda et al. 2016; Gregor et al. 2018). Owing to its high luminescent activity and small size (19 kDa),

NanoLuc™ luciferase generated from the deep sea shrimp *Oplophorus gracilirostris* has recently become a favoured bioluminescent reporter and been used in whole-cell biosensors (Mary P Hall et al. 2012; Cevenini et al. 2018). Interestingly, many of these biosensors with bioluminescent outputs have been coupled with portable devices for field testing or on-site diagnosis (Roda et al. 2011; Siegfried et al. 2012; Cevenini et al. 2018; Mimee et al. 2018), such as the ARSOlux kit which is currently tested in six countries for arsenic contamination (http://www.ufz.de/arsolux/index.php?en=20706). However, unless the whole luciferase cassette (e.g., *luxCDABE*) is present, external substrate such as D-luciferin, coelenterazine or furimazine will be required to generate luminescence, limiting their applications for continuous monitoring.

1.2.2.2 Fluorescent output

Compared to luciferase, proteins that can produce colour or light without substrates are more convenient for continuous measurement. Fluorescent proteins can be used as they are relatively stable, taking a short time to mature, and their emission light can be easily measured by a fluorimeter with specific light excitation. Some enzymatic reactions also generate fluorescence, such as β-galactosidase (LacZ) with fluorescein di-beta-D-galactopyranoside (FDG)-based substrates (Rowland et al. 1999; Didovyk et al. 2017). The fluorescence is also visible by eye under a light with suitable excitation wavelength. Green fluorescent protein (GFP) from the jellyfish *Aequeora victoria* is probably the most frequently used fluorescent protein for biosensing (Shimomura et al. 1962; Buffi et al. 2011; Saeidi et al. 2011; Wang et al. 2013; Hwang et al. 2014; Cerminati et al. 2015; Courbet et al. 2015; Kim et al. 2016; Merulla and van der Meer 2016; Daeffler et al. 2017; Kim et al. 2018). GFP mutants derived from wild type GFP have been developed, such as GFPmut3 for better excitation at 488 nm, superfolder GFP for improved fluorescence intensity (Pédelacq et al. 2006), superfast GFP for fast protein folding, and deGFP for better translation in cell-free systems (Shin and Noireaux 2010). Other fluorescent proteins with different excitation and emission wavelengths have also been identified or developed (e.g., mCherry derived from DsRed found in *Discosoma striata*), which can be selectively used to avoid cross-talk with auto-fluorescence or fluorescence background from cells (Shaner et al. 2004; Subach et al. 2011). However, external excitation light and the emission light reader are required for the measurement, which may not be convenient to use in the field.

1.2.2.3 Electronic output

Electronic signals have been used as a sensor output which can be measured by electrodes (Webster et al. 2014; Pous et al. 2018). Webster et al. (2014) developed a novel

bioelectrochemical system in engineered bacteria to sense arsenic and report it via an electric signal, which can be convenient for commercial remote monitoring. *Shewanella oneidensis*, a natural heavy metal reducing bacterium, was used as the chassis for building the arsenic sensing and anode reducing circuits. The Δ*mtrB S. oneidensis* was reconstructed with an ArsR-P_{arsR}-*mtrB* device, which drove MrtB expression once sensing the As^{3+}. MrtB is indispensable for the electrode reduction. This bioelectrochemical system can be modified in a modular way by replacing the metal sensing device, so that the system would be able to sense and report other input signals.

1.2.2.4 Colorimetric output

Colorimetric reporters are often used to allow direct visualisation of sensor output by the naked eye. These include some fluorescent proteins, chromoproteins and enzymes that produce pigments or other coloured products (Biran et al. 2003; Stocker et al. 2003; Fujimoto et al. 2006; Wackwitz et al. 2008; Joshi et al. 2009; De Mora et al. 2011; Joe et al. 2012; Shin 2012; Kotula et al. 2014; Pardee et al. 2014; Courbet et al. 2015; Danino et al. 2015; Huang et al. 2015a; Pardee et al. 2016a; Didovyk et al. 2017; Watstein and Styczynski 2018). Fluorescent proteins and chromoproteins will need sufficient accumulation to generate visible colours, which may reduce sensor sensitivity and increase response time. Enzyme-based colour change is generally fast but may require additional substrates.

LacZ from *E. coli*'s *lac* operon is probably the most popular enzyme used for bacterial sensors (Biran et al. 2003; Stocker et al. 2003; Wackwitz et al. 2008; Joshi et al. 2009; De Mora et al. 2011; Joe et al. 2012; Shin 2012; Kotula et al. 2014; Pardee et al. 2014; Danino et al. 2015; Huang et al. 2015a; Pardee et al. 2016a; Didovyk et al. 2017). Different substrates have been used for LacZ-based colorimetric output, such as colourless X-gal (i.e., 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside) which can be finally converted to an insoluble blue compound 5,5'-dibromo-4,4'-dichloro-indigo, and yellow CPRG (i.e., chlorophenol red-β-D-galactopyranoside) which can be converted to a purple chlorophenol red product. The LacZ catalysed hydrolysis of these substrates is generally fast; moreover, CPRG is reported to be more sensitive than X-gal and the colour change can be visualised in 1 h (Möckli and Auerbach 2004; Pardee et al. 2014). However, many bacteria naturally contain a *lac* operon, which will increase the background of LacZ. Apart from deleting or inactivating the *lacZ* in the bacteria, other enzyme-based colorimetric outputs can be considered, such as chitinases which can cleave a colourless substrate, 4-Nitrophenyl N,N'-diacetyl-β-D-chitobioside, and yield a yellow p-nitrophenol product (Pardee et al. 2014). Nevertheless, the CPRG and 4-Nitrophenyl N,N'-diacetyl-β-D-chitobioside are relatively expensive, and require careful storage at low

temperature, which may increase the difficulty and cost of using these reporters for large scale sample screening and field testing.

To reduce the cost of using LacZ as a reporter, a unique strategy based on pH change has been developed in *E. coli* for testing arsenic (Joshi et al. 2009; De Mora et al. 2011). This strategy is based on the fact that LacZ cleaving lactose allows fermentation in *E. coli* and consequent acid production. Bromthymol blue (yellow in acidic solutions, blue in basic solutions) has been used as a pH indicator to monitor this process, and the colour change can be visualised by the naked eye after 24 h. Other reporters can also be used to reduce the cost of substrates, such as cytochrome c peroxidase and XylE (catechol 2,3-dioxygenase) (Wackwitz et al. 2008; Joshi et al. 2009). Their substrates guaiacol and catechol are relatively cheap, and they can develop visible brown and yellow colours respectively; however, both guaiacol and catechol are not stable and need to be stored under restricted conditions.

Pigments such as carotenoids and violacein have been used as output in biosensors (Fujimoto et al. 2006; Joe et al. 2012; Watstein and Styczynski 2018). They can be produced without substrates but require sophisticated metabolic pathways which may slow down the sensing response. To avoid engineering those pathways and using extra substrates, specific cell strains (e.g., *Rhodovulum sulfidophilum* and *Deinococcus radiodurans*) that normally contain those pathways can be used to make biosensors (Fujimoto et al. 2006; Joe et al. 2012). In this case, only the enzyme that catalyses the final colour change will be needed as reporter, such as CrtA (i.e., spheroidene monooxygenase) which produces red spheroidenone from yellow spheroidene (Fujimoto et al. 2006), and CrtI (i.e., phytoene dehydrogenase) which converts colourless phytoene to red lycopene (Joe et al. 2012).

1.2.2.5 Output for in vivo monitoring

There are also reporters that enable the monitoring of biosensors inside animals in real time, such as a recently developed acoustic reporter and the aforementioned luciferase. The acoustic reporter can generate intracellular gas vesicles in *E. coli*, which can be detected by general ultrasound techniques (Bourdeau et al. 2018). Biosensors with luciferase as reporter have been entrapped in an ingestible micro-bio-electronic device (IMBED) that can detect analytes which can diffuse into it; the emitted light can be converted into digital signals through photodetectors and the signals can be further transmitted to computers via Wi-Fi (Mimee et al. 2018).

1.2.2.6 Alternative functional outputs

Alternative functional outputs have also been used for biosensors, and endow them with specific abilities upon sensing, such as moving to a specific target and killing targeted cells. Hwang et al. (2014b) used bacterial chemotaxis mechanism along with killing systems to seek and kill biofilm-encased pathogenic *P. aeruginosa*. Three output genes were fused with *P. aeruginosa*'s quorum sensing system (i.e., the LasR-AHL system) in *E. coli*: **1)** *cheZ*, which facilitated bacterial movement to pathogens upon sensing AHL, **2)** nuclease gene *DNaseI* which helped in degrading the pathogens' biofilm, and **3)** *mcsS* which expressed an antimicrobial peptide, microcin S, to kill the pathogen. Other bacteriocins have also been used as outputs to kill *P. aeruginosa*, such as pyocin S5 and a more specific chimeric bacteriocin CoPy (Saeidi et al. 2011; Gupta et al. 2013).

Various outputs have been developed for bacteria-based cancer therapy. For example, an invasion gene from *Yersinia pseudotuburculosis* was fused to an anaerobically induced *fdhF* promoter in *E. coli*; when they sense a hypoxic environment (e.g., a tumour micro-environment), the invasin will be expressed and allow *E. coli* to penetrate into mammalian cells for cancer therapy (Anderson et al. 2006). Apart from cell invasion, a plant enzyme, myrosinase, has been used as an output to convert dietary glucosinolate to sulforaphane, which can inhibit cancer cell growth and promote apoptosis (Ho et al. 2018).

Another interesting functional reporter is ice nucleation protein (e.g., InaZ from *Pseudomonas syringae pv. syringae*) which can promote ice nucleation at relative high temperature. Although the measurement of ice nucleation may not be as easy as the aforementioned reporters, it has been shown that such reporter provides better output dynamic ranges at low target concentrations when compared to LacZ and GFP (Miller et al. 2001).

1.2.3 Chassis choice

Generally, a synthetic biosensor can be built in two ways: **(1)** using a host endogenous genetic pathway and rewiring its final output to a desired reporter gene (**Fig. 1.1d**), or **(2)** importing a heterogeneous signalling pathway from a specialised species to another species (**Fig. 1.1e**) (Bernard and Wang 2017).

The first way is simple and the sensor is likely to work as all the relevant components in the pathway are in the host. However, this sensor only senses its target with one particular sensitivity, selectivity and output dynamic range, which may not be optimal for real applications; also, the sensing pathway may be crosslinked to other components in the host, so the sensing response may vary under different environmental conditions (Bernard and

Wang 2017). Despite the sensing variation caused by crosstalk, sensor sensitivity, selectivity and output dynamic range can be regulated using synthetic biology tools, which will be described in Section 1.3.

The second way is usually applied when the native bacterial sensors are not suitable for study or application due to their difficult/incompatible growth condition or pathogenic features. As the most studied prokaryotic model organism, *E. coli* has been widely used for studying and engineering biosensors. It is easy and cheap to culture and genetically manipulate, and it has been endowed with various sensing abilities originating from other bacteria. Moreover, *E. coli* Nissle 1917, a probiotic strain, has been particularly applied for studying *in vivo* diagnosis and biotherapy (Duan and March 2010; Gronbach et al. 2010; Danino et al. 2015; Daeffler et al. 2017; Ho et al. 2018). *Bacillus subtilis* is also a popular chassis for biosensing as it can form spores for long term storage (Date et al. 2007; Joshi et al. 2009; Courbet et al. 2015; Volpetti et al. 2017). Notably, using heterogeneous components may eliminate crosstalk between the sensing pathway and endogenous machinery from the host (Wang et al. 2011a; Liu et al. 2018a). However, the imported pathway's compatibility with the new host is not guaranteed, and neither its functionality. Therefore, each component from the heterogeneous pathway needs to be characterised and optimised carefully in the new host (Bradley et al. 2016b; Bernard and Wang 2017).

1.2.4 State-of-the-art of synthetic biology in biosensing

The advantages of synthetic biology can be summarised in four principles: **1)** Abstraction: an abstractive hierarchy is used to manage complexity while less-related information is hidden for clarity. **2)** Standardisation: genetic parts are standardised into reusable parts with defined parameters measurement. **3)** Modularisation: genetic parts or devices are independent modules with defined functions and are interchangeable. **4)** Rational and quantitative design: the behaviour of genetic parts or devices can be predicted by mathematical models. These advantages allow synthetic biology communities to develop diverse collections of sensing promoters and reporters or actuators, and also to develop new genetic devices for building novel biosensors.

With the emerging techniques of genome sequencing, DNA assembly and other molecular biology techniques, it is now easy to identify and test promoters, their regulators and other functional genes. With the rational design strategies from synthetic biology, the modular and standardised parts can be easily used to make biosensors in a 'plug-and-play' fashion (**Fig. 1.1a**).

1.3 Tools and strategies from synthetic biology for improving biosensors

Prokaryotic biosensors can be constructed easily by simply combining the sensing modules and output modules. However, as their elements are from nature, those modules may have certain limitations when it comes to real world applications, especially for selectively and sensitively sensing particular targets. For example, some naturally occurring heavy metal responsive receptors can respond to multiple ions, which will reduce their sensing specificity in a complex environment (Amaro et al. 2011; Wang et al. 2013). For example, CadC from *S. aureus* can sense Cd^{2+}, Pb^{2+} and Zn^{2+}, and CmtR from *Mycobacterium tuberculosis* can sense Cd^{2+} and Pb^{2+} (Saha et al. 2017). Additionally, some native promoters are relatively weak, which may limit the sensor's sensitivity and output detection (Stocker et al. 2003; De Mora et al. 2011; Siegfried et al. 2012; Wang et al. 2013; Huang et al. 2015a; Kim et al. 2016; Merulla and van der Meer 2016).

Fortunately, the emerging discipline of synthetic biology has developed numerous tools and strategies which can be used to improve and customise sensor performance. Mathematical models have also been developed to facilitate sensor improvement in a quantitative manner (Ang et al. 2013; Berset et al. 2017; Mannan et al. 2017), but this section only focuses on tools and strategies that have been experimentally demonstrated in prokaryotic biosensors.

1.3.1 Properties of a biosensor

From the perspective of engineering, it is important to define measurable properties of a biosensor so that each property can be customised quantitatively to improve the final sensor performance. Different metrics have been used to define sensor performance (Mannan et al. 2017). Here, I focus on four properties which are closely related to sensor application: selectivity, sensitivity, output dynamic range and leakiness (**Fig. 1.2**). Response time is another important parameter for characterising sensors' performance but is not discussed here.

Selectivity is a qualitative property, meaning how well a sensor can distinguish a target of interest among other similar molecules or signals (**Fig. 1.2a**). The other three properties can be quantitatively measured and optimised. Most biosensor response curves are sigmoidal, where the output increases with the target concentration. In this case, the sensitivity can be defined as limit of detection (LOD) which is the minimal target concentration that causes significant output increase of a sensor. Output dynamic range can be defined as the ratio of the maximum output level (achieved by target induction) to the basal output level (when no target is present). The basal output level is considered as leakiness, and high leakiness may cause the output to be not distinguishable between sensors with or without target. The synthetic biology-related solutions to improve each of the four properties are described in the following sections.

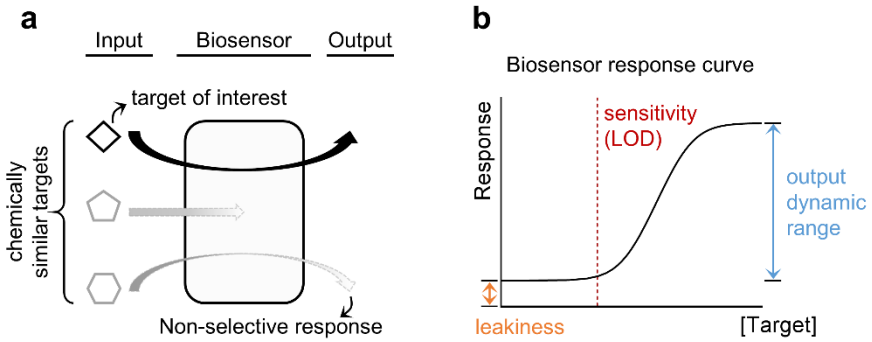


Figure 1.2: Metrics for defining performance of a biosensor.

(**a**) Schematic illustrating a biosensor's selectivity. (**b**) A biosensor response curve with sensor's leakiness, sensitivity and output dynamic range annotated. The sensitivity can be defined as the limit of detection (LOD).

1.3.2 Strategies for improving selectivity

Generally, the specificity of receptors can be improved by altering amino acid residues of their effector recognition domain. For example, a modified CueR lost the ability to respond to Ag^{+} but became more sensitive to Au^{+} (Stoyanov and Browns 2003), a mutated MerR improved its ability to sense Cd^{2+} rather than Hg^{2+}, Zn^{2+} and other tested metals or metalloids (Hakkila et al. 2011), and a RcnR mutant specifically responded to Ni^{2+} rather than Co^{2+} (Cayron et al. 2017). Such random mutagenesis can be extremely helpful especially when relevant sensing and responding elements are hidden. However, despite notable successes, the mutation is not predicable, and it is time consuming to find desired mutations. By contrast, the rational design from synthetic biology can provide predictable solutions based on available information.

Logic gates, such as AND, NAND, NOR and XOR gates, have been widely used to build complex genetic circuits as they can combine multiple sensing and output modules. The AND gate is one of the simplest gates; it only operates in the presents of two inputs. Due to this property, it can be used in biosensors to only express output when simultaneously sensing multiple inputs. This method has been used to improve sensor specificity by filtering out the response to non-relevant targets (Wang et al. 2013). In this case, non-specific receptor proteins ZraR (responding to Zn^{2+} and Pb^{2+}) and ZntR (targeting Zn^{2+} and Cd^{2+}) were combined in an AND gate, resulting in a sensor selectively responding to Zn^{2+} (**Fig. 1.3a**).

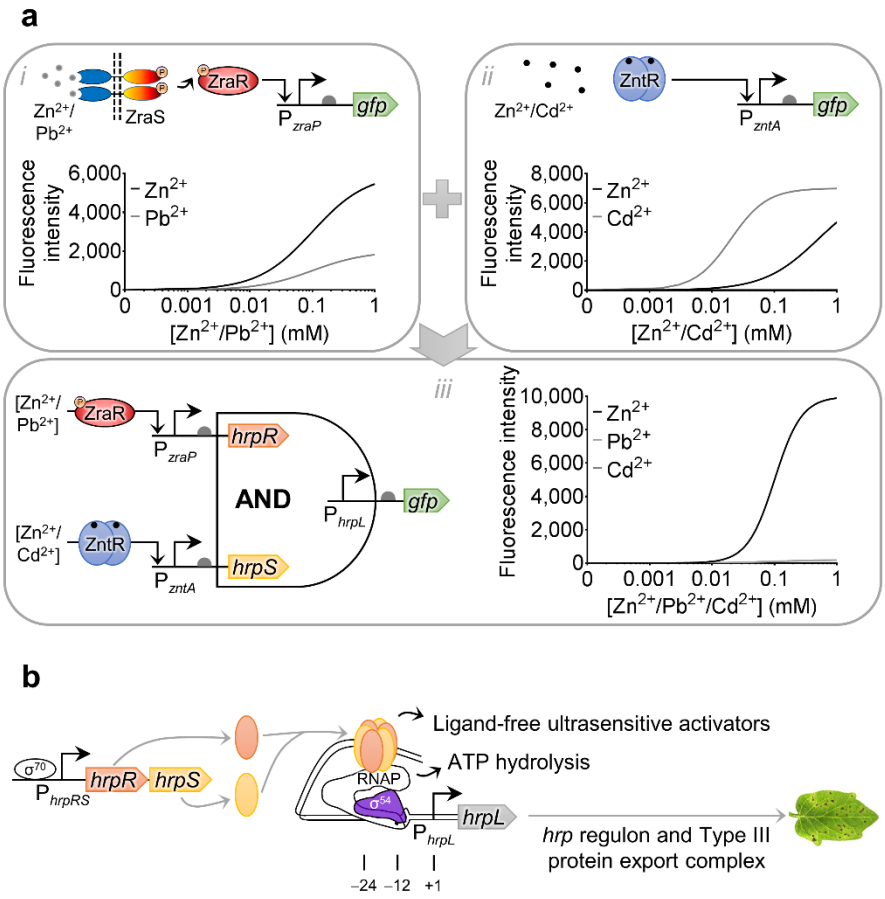


Figure 1.3: Biosensor specificity enabled by synthetic biology.

(**a**) A zinc-specific biosensor using an AND gate (Wang et al. 2013). (**a-*i***) A zinc/lead biosensor and its response curves for Zn^{2+} and Pb^{2+}. (**a-*ii***) A zinc/cadmium biosensor and its response curves for Zn^{2+} and Cd^{2+}. (**a-*iii***) A zinc-specific biosensor was generated by integrating both sensing modules from ***i*** and ***ii*** into an AND gate. *gfp*, gene encoding a green fluorescent protein. (**b**) The HrpR/HrpS hetero regulation motif in the *hrp* (hypersensitive response and pathogenicity) system of *P. syringae* pv. tomato DC3000 (Wang et al. 2011a; Wang and Buck 2014). The *hrp* system promotes pathogenicity of the bacterium in its plant host. The σ^{54}-dependent *hrpL* promoter can be activated by the heterohexamers of the transcriptional activators HrpR and HrpS.

The AND gate mentioned above is based on the HrpRS activator complex and its cognate promoter P_{hrpL} (**Fig. 1.3b**). This system is originally from the *hrp* (hypersensitive response and pathogenicity) gene cluster of a plant pathogen, *P. syringae*. P_{hrpL} is a σ^{54}-dependent promoter, which can be activated only when both HrpR and HrpS are present and form a hetero-hexamer on the promoter (Jovanovic et al. 2011; Wang et al. 2011a). In addition to the HrpRS system, other systems have been used for making AND gates in bacterial sensors, such as a T7 RNAP-amber suppressor system to integrate 2 input signals (Anderson et al. 2007), activator-

chaperone systems from *Salmonella*, *Shigella* and *Pseudomonas* used to make a 4-input AND gate (Moon et al. 2012), and split T7 PRNP systems (Shis and Bennett 2013; Schaerli et al. 2014). Integrases have also been used to make AND gates, for example, by using them to flip the direction of genetic parts (e.g., terminator, promoter and gene); only when all the genetic parts are in the right direction will the final output be expressed (Bonnet et al. 2013; Courbet et al. 2015). Integrase function is described further in Section 1.4.1.2.

1.3.3 Strategies for enhancing sensitivity

Similar to the traditional solutions for biosensor selectivity, random mutagenesis of receptors and promoters is a popular way to improve sensor sensitivity (Li et al. 2015; Daeffler et al. 2017). However, as the sensitivity is intrinsically related to the relative concentrations of the receptor and its ligand, tuning the receptor and ligand intracellular densities should be a more rational and predictable approach.

1.3.3.1 Sensitivity improvement by tuning receptor density

Prokaryotic biosensor sensitivity can be regulated by logic operations. A basic prokaryotic biosensor can be considered as a buffer gate: the output will only be on when there is an input signal. The buffer gate can be a repressor or an activator-based system (**Fig. 1.4a,b**). For example, the ArsR-P_{arsR} arsenic sensor is based on ArsR repressor, whereas the LuxR-P_{lux} AHL sensor is based on LuxR activator (**Fig. 1.4c,d**). In both cases, when the targets (i.e., As^{3+} or $3OC_6HSL$) are present in sufficient concentrations, the sensors will be on (**Fig. 1.4a,b**). The repressor and activator can be expressed from a constitutive promoter (P_C); the sensor's sensitivity to their targets can be easily tuned by the expression level of their repressor/activator, e.g., by using weaker or stronger P_C (Merulla et al. 2013; Wang et al. 2015). In detail, with a weaker P_C for ArsR expression, fewer protein molecules of ArsR will be present in cells and fewer will be free from the P_{arsR}, therefore the As^{3+} can more easily find the P_{arsR}-bound ArsR molecules and release them from the promoter; on the contrary, if more ArsR molecules are expressed, more of them will be free in the cytoplasm and available to bind to As^{3+}, so less P_{arsR}-bound ArsR will be released from P_{arsR} and less P_{arsR} will be activated (**Fig. 1.4c**). However, there is a limit in ArsR tuning as no or very few ArsR will lead to high leakiness and no sensing response (Wang et al. 2015). For an activator-based system like the LuxR-P_{lux} AHL sensor, more LuxR present in the cells will increase the chance of forming LuxR-AHL complex, therefore more P_{lux} can be activated by the complex and the sensor is more sensitive (**Fig. 1.4d**). Moreover, the output dynamic will also be increased by tuning the P_C (Wang et al. 2015). Overall, by tuning the regulator intracellular densities, both sensor sensitivity and output dynamic range can be improved.

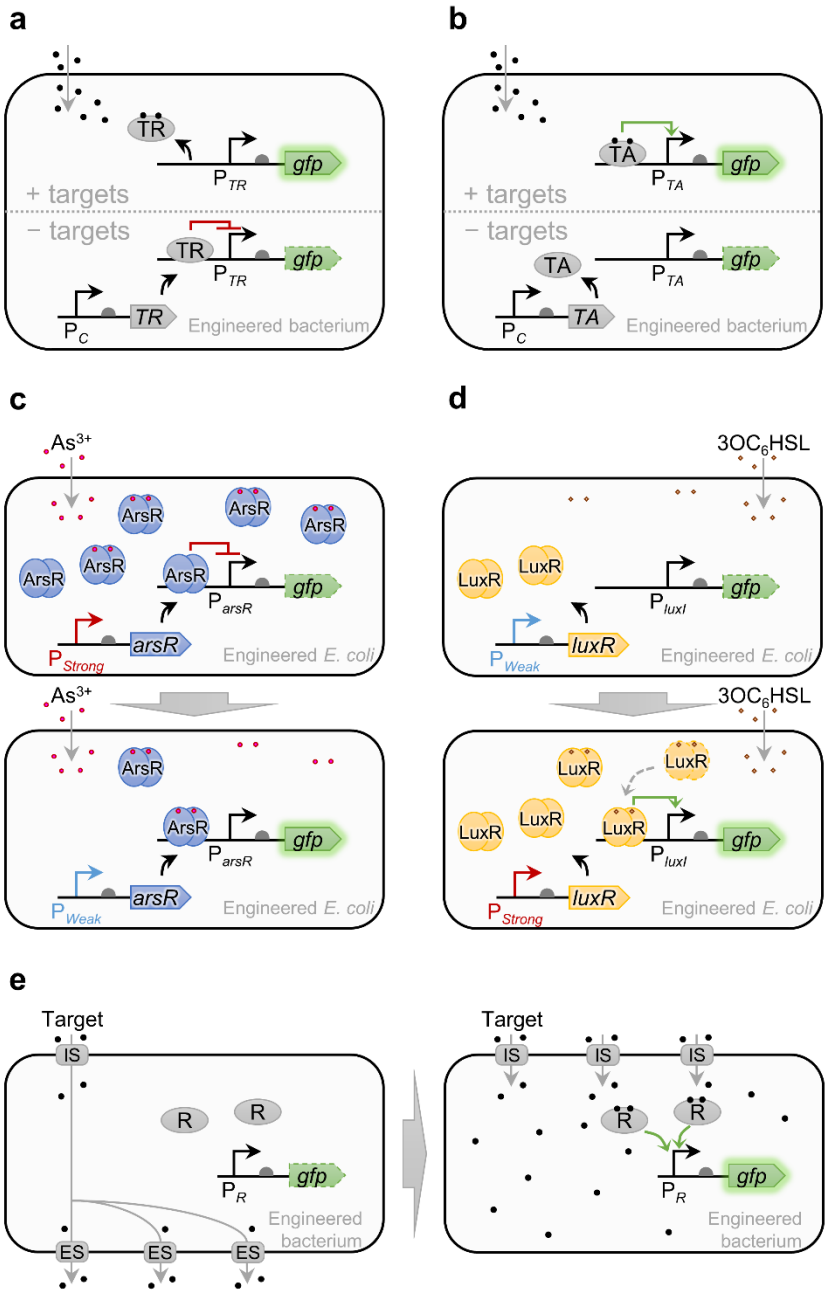


Figure 1.4: Strategies for improving biosensor sensitivity.

(**a** and **b**) A transcriptional repressor (TR) and a transcriptional activator (TA)-based biosensor. P_C, constitutive promoter. P_{TR}, TR's cognate promoter. P_{TA}, TA's cognate promoter. Black dots, targets of interest. *gfp*, gene encoding a green fluorescent protein. (**c** and **d**) Improving a biosensor's sensitivity by tuning receptor densities (Wang et al. 2015). **c** shows a TR-based arsenic biosensor, and **d** shows a TA-based AHL biosensor. P_{Strong}, strong constitutive promoter. P_{Weak}, weak constitutive promoter. (**e**) Improving a biosensor's sensitivity by increasing its targets' intracellular density (Hynninen et al. 2010; Cayron et al. 2017). IS, import system. ES, export system. R, receptor. P_R, R's cognate promoter.

A toggle-switch has been used for improving a cadmium sensor sensitivity dynamically (Wu et al. 2009). In this system, an IPTG inducible promoter P_{tac} controls the expression of the cadmium receptor CadR; a cadmium inducible promoter P_{cadR} controls the expression of LacI and GFP. P_{tac} can be repressed by LacI unless there is IPTG, whereas P_{cadR} can be repressed by CadR unless there is cadmium. This toggle switch design can lead to a more robust sensing ability due to positive feedback. However, the real reason that makes the system more sensitive is still the regulation of the receptor. With low IPTG induction, the CadR intracellular density can be low, and therefore it increases the sensitivity; with cadmium induction, CadR intracellular density will be even lower, which further improve the sensitivity and the output dynamic range. Therefore, the amount of IPTG is the key to the sensor sensitivity.

1.3.3.2 Sensitivity improvement by tuning intracellular ligand density

Another option to improve sensor sensitivity is to improve target intracellular density by increasing the levels of its import system while deleting the export system in host cells (**Fig. 1.4e**). For example, by disrupting efflux transporters for Zn/Cd/Pb in *P. putida* KT2440, the bacterial detection limits against Zn/Cd/Pb were improved up to 45-fold (Hynninen et al. 2010); by introducing several Ni-uptake systems and deleting the Ni efflux pump, an engineered *E. coli* biosensor achieved much better sensitivity and output dynamic range (Cayron et al. 2017). However, this method may not be applicable to some sensing systems where no specific import or export systems are required for the targets.

1.3.4 Strategies for improving output dynamic range

A number of native promoters are relatively weak and thus their derived biosensors have limited output dynamic ranges (Stocker et al. 2003; De Mora et al. 2011; Siegfried et al. 2012; Wang et al. 2013; Huang et al. 2015a; Kim et al. 2016; Merulla and van der Meer 2016). Again, random mutagenesis of promoters is often performed to maximise transcriptional output.

1.3.4.1 Improving output dynamic range by tuning σ binding sites

Recently, however, a more rational mutagenesis procedure has been proposed to regulate promoter output dynamics (Chen et al. 2018). The principle of this regulation is to change the binding affinity of the promoter to its σ factor/RNAP; the mutation mainly occurs at the sigma factor binding sites (e.g., –10 and –35 regions for σ^{70}) (Cox et al. 2007; Brewster et al. 2012; Guzina and Djordjevic 2017; Chen et al. 2018). Although there are no clear rules for the site mutation, Chen et al. (2018) recently developed and tested a small library of –10 and –35 regions with defined input and output dynamics, which can be easily applied to σ^{70}-based sensing systems (**Fig. 1.5a**). However, this library is only useful for σ^{70}-based promoters, and

it may be difficult to regulate some sensing systems where the operator naturally overlaps with σ binding sites.

1.3.4.2 Improving output dynamic range through transcriptional amplifiers

A more universal and predictable way to improve sensor output dynamic range is to use transcriptional amplifiers (TAmps). A TAmp is like an electronic amplifier, and it amplifies a transcriptional signal before outputting it. A TAmp is basically a one-input logic gate (i.e., NOT or Buffer gate), and the input is the output from the sensing module (i.e., the expression from the sensing promoter like P_R) (**Fig. 1.5b**). Generally, a TAmp comprises a transcription factor (TF) and its cognate promoter P_{TF}; the TF is under the control of a sensing promoter, and the P_{TF} controls the expression of an output gene. More importantly, to qualify as an amplifier, P_{TF} must have higher maximum activity than P_R, so that a small transcriptional input from P_R can be converted into a larger downstream output.

Early examples of TAmps were constructed based on transcriptional repressors, which were cascaded to an AHL sensor and shown to improve sensor output dynamic range (Karig and Weiss 2005). The same team also built 2 and 3-layered repressor cascades, which were shown to improve the output dynamic range of an anhydrotetracycline (aTc) sensor (**Fig. 1.5b-*ii***) (Hooshangi et al. 2005). However, unless the repressors were cascaded in even numbers, repressor-based amplifiers (TR-Amp) may not be suitable for sensing modules that have positive relationships between input and output. A positive relationship means that the output will be on when there is an input.

TAmps based on transcriptional activators also have been developed (**Fig. 1.5b-*iii***). The activator-based amplifier (TA-Amp) has been exemplified by a TAmp constructed from the aforementioned HrpRS system (Wang et al. 2014). This amplifier was demonstrated to significantly improve an arsenic sensor's output dynamic range, and it amplified the output in a linear manner within a wide input range. Moreover, the amplification ability can be tuned by regulating the level of HrpS, either at translational (i.e., tuning the strength of ribosome binding site (RBS)) or post-translational level (**Fig. 1.5c**) (i.e., tuning the expression of a HrpS inhibitor HrpV). Another recent example using T7 RNAP and P_{T7} as a TA-Amp was demonstrated to improve a cadmium/lead sensor's sensitivity and output dynamic range (Kim et al. 2016). Unlike the TR-Amp, only one layer of TA-Amp is needed for amplifying sensing modules with positive relationships between input and output.

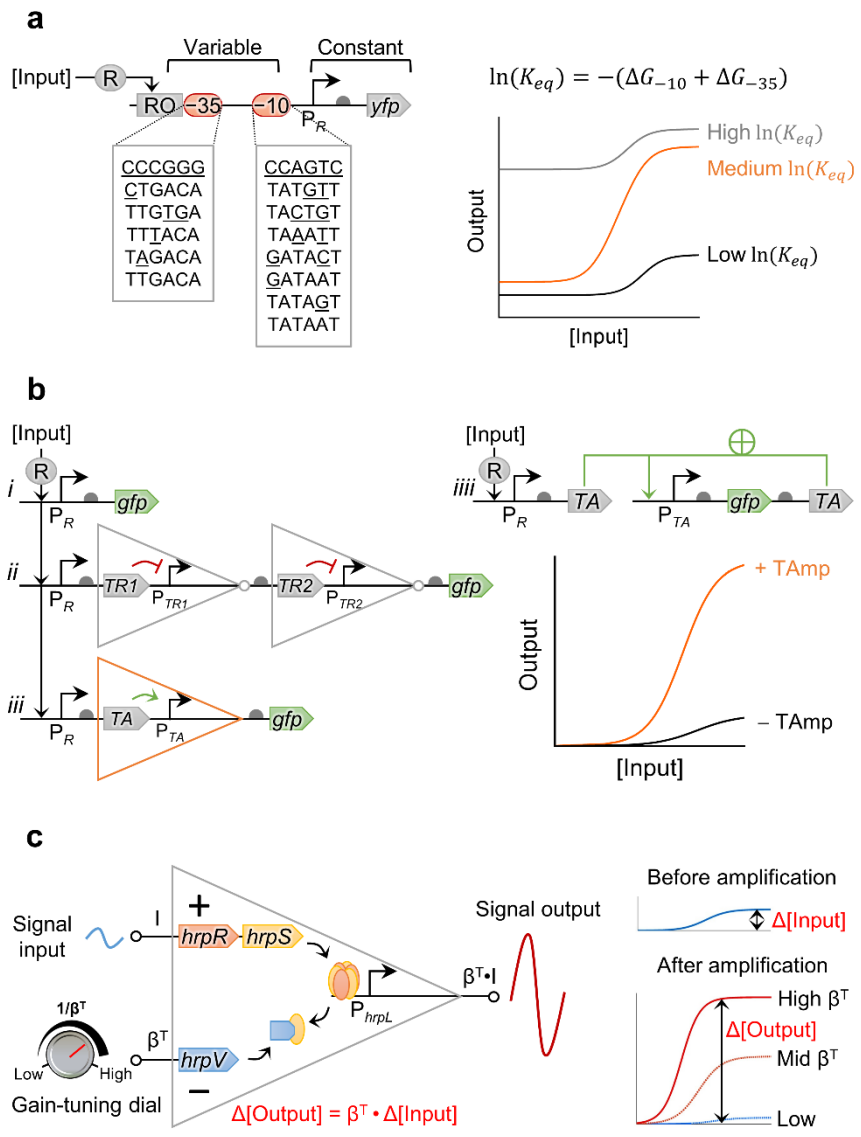


Figure 1.5: Strategies for improving biosensor output dynamic range.

(**a**) Different σ^{70} binding sites of an inducible promoter yield different basal levels and output dynamics as a result of the relative equilibrium constants of σ^{70} binding to the −10 and −35 regions, $\ln(K_{eq}) = -(\Delta G_{-10} + \Delta G_{-35})$ (Chen et al. 2018). Non-consensus bases are underlined. ΔG_{-10} and ΔG_{-35} correspond to the relative changes in the binding energy due to changes in the −10 and −35 sites. R, receptor. P_R, R's cognate promoter which contains an operator for R (RO). *yfp*, gene encoding a yellow fluorescent protein. (**b**) Biosensors without transcriptional amplifiers (TAmp) (**b-*i***), with two transcriptional repressor-based amplifiers (**b-*ii***), or with a transcriptional activator-based amplifier (**b-*iii***), and their response curves (Hooshangi et al. 2005; Wang et al. 2014; Kim et al. 2016). ***iiii***. An amplifier with positive feedback (Nistala et al. 2010). P_{TR}, TR's cognate promoter. P_{TA}, TA's cognate promoter. *gfp*, gene encoding a green fluorescent protein. (**c**) A gain-tuneable TAmp based on a HrpRSV system (Wang et al. 2014). This device scales the weak transcriptional input signal (I) linearly in response to a second 'gain tuning' transcriptional input (β^T).

If the sensing module is tightly regulated, the amplification capacity of a TAmp can be further enhanced through a positive feedback loop (**Fig. 1.5b-*iiii***). For example, Nistala et al. (2010) built a TA-Amp based on the LuxR-P_{lux} activation system, where a mutated LuxR was used which activated P_{lux} without extra inducers. Instead of fusing the P_{lux} with a *gfp* output alone, *luxR* was also fused within the same operon; therefore, when the system is activated, more LuxR will be produced to further activate the system. Compared to the TA-Amp alone, this TA-Amp coupled with positive feedback further and significantly improved the output of an aTc sensor and an aspartate sensor. Notably, it also improved the detection limit of the aTc sensor.

Output amplification can also be performed using recombinase-based memory modules (Bonnet et al. 2013). In this case, the final output expression can be activated by the recombinase, and the maximal transcriptional activity must surpass that of the sensing promoter. Details will be described in Section 1.4.1.2.

1.3.5 Strategies for reducing leakiness

In addition to low sensitivity and low output, high leakiness is also a common issue for biosensors. When a sensor is leaky, it expresses obvious output even in the absence of its target. Although the leakiness can be ignored under some circumstances, it may cause issues in many other cases: **1)** Enzyme-based colorimetric output may easily saturate with high basal level production, which will restrict titrimetric analysis (Wackwitz et al. 2008). **2)** It may affect the downstream regulatory elements and shut down the dynamic output (Nielsen et al. 2016). **3)** It may cause protein overexpression, which can be toxic to host cells, reduce protein functions (Tan et al. 2013; Kuznetsova et al. 2014), affect resource allocation (Gyorgy et al. 2015), and thereby interfere with the response of genetic circuits (Qian et al. 2017; Venturelli et al. 2017). **4)** It may cause health problems if the sensor outputs have killing effect for biotherapy.

Most sensing systems from nature display more or less leakiness, and we cannot always find a less leaky sensing element with the desired functions from nature. Therefore, many studies have been carried out to reduce leakiness. Here, I classify these methods into three types, where the leakiness are managed at **1)** transcriptional level, **2)** translational level or **3)** post-translational level (**Fig. 1.6**).

1.3.5.1 Managing leakiness at transcriptional level

One traditional way to reduce the leakiness of a sensing system is to randomly mutate the sensing regulators and promoters in the hope of generating forms with less leakiness while maintaining the sensing abilities. This mutation basically regulates the transcription initiation

by changing the affinity between the regulators and the promoters. One study combined error-prone PCR (i.e., polymerase chain reaction) and FACS (i.e., fluorescence-activated cell sorting) to mutate and sort arsenic inducible operons for low leakiness and high output dynamic range (Li et al. 2015). The selected mutations were at the ArsR binding site (ABS) in P_{arsR} and the DNA binding site in ArsR, which may have improved the affinity between ArsR and ABS, and therefore significantly decreased sensor leakiness. Similar methods can also be applied to other regulators and promoters. However, this random mutation and screening is costly and time consuming.

Since the sensing promoter is the major source of leakiness, it is more logical to directly modify the leaky promoter. The aforementioned mutagenesis at σ^{70} binding sites can also reduce sensor leakiness (**Fig. 1.5a**) (Chen et al. 2018). A more straightforward and predictable approach is to engineer the structure of leaky promoters, such as by **1)** regulating numbers and positions of an operator in a promoter (**Fig. 1.6a,b**) (Merulla and van der Meer 2016; Chen et al. 2018), and **2)** adding an antisense promoter to the leaky sensing promoter (**Fig. 1.6c**) (Brophy and Voigt 2016).

Regulation through operators

The operator is a part of a promoter sequence, where a regulator can bind and control the activation of the promoter. It has been shown that an activator's (e.g., AraC and LuxR) operator can only work upstream of the –35 region in the promoter (i.e., distal region), whereas a repressor's (e.g., LacI and TetR) operator can work not only in the distal region, but also between –10 and –35 regions (i.e., core region) and downstream of the –10 region (i.e., proximal region). Moreover, it has been suggested that the repression efficiency at those places is core ≥ proximal ≥ distal (**Fig. 1.6a**) (Cox et al. 2007). In addition, if the repressor's operator shares the sequence of the –10 and –35 regions, it can further improve the repression efficiency and therefore reduce the leakiness of the promoter; also, if a ligand-responsive activator (e.g., AraC and RhlR) can bind to its operator at the distal region in the absence of the target ligand, it can make a promoter tighter (i.e., with lower basal expression) (Chen et al. 2018).

Additionally, if the target's regulator is a repressor, adding an extra operator for this repressor downstream of a promoter can further reduce the basal expression level of the promoter; this phenomenon is based on 'roadblocking' where the extra repressor binding downstream of the promoter inhibits the read-through of the RNAP (**Fig. 1.6b**). Also, tuning the distance between the extra operator and the core region can regulate the roadblocking efficiency thus regulating promoter basal expression (Hao et al. 2014; Merulla and van der Meer 2016). However, the roadblocking effect can be affected by promoter strength, regulator concentration and the

affinity of the regulator to its operator (Hao et al. 2014). Therefore, this approach may work differently under different circumstances, and detailed characterisation will be required for different promoter-operator systems.

Regulation through antisense transcription

Antisense transcription arises from promoters, and the antisense promoter drives RNAP reading in the opposite direction but towards a target promoter. It was first discovered in bacteria more than 40 years ago (Pelechano and Steinmetz 2013). In terms of leakiness reduction, the most important function of antisense promoters is to interrupt the RNAP reading through from the target promoter (**Fig. 1.6c-*iii***) (Brophy and Voigt 2016), thus interrupting leaky transcription and reducing background expression. Based on the location of the antisense promoter, it may also regulate the leakiness by interrupting translation: **1)** if the antisense promoter is located a short distance downstream of its target promoter, it may produce short antisense RNA to inhibit the translation initiation of the target gene (**Fig. 1.6c-*i***) (Kawano et al. 2007); **2)** if the antisense promoter is located at the 3' end of the target promoter and gene, it may produce long antisense RNA to form double-strand RNA with target gene's mRNA and trigger RNA degradation (**Fig. 1.6c-*ii***) (Lasa et al. 2011). Notably, the second approach will be easier to use in terms of sequence design, and the target promoter activity can be easily tuned by regulating the strength of the antisense promoter (**Fig. 1.6c**) (Brophy and Voigt 2016).

1.3.5.2 Managing leakiness at translational level

The leaky expression happens at both transcriptional and translational levels; thus reducing the translation efficiency is another way of reducing leakiness, and it can be very helpful if the transcription rate cannot be easily changed. Similar to the regulator/promoter mutagenesis, mutations can also be introduced to the RBS; changing the RBS sequence changes its affinity for ribosome binding and therefore regulates the initiation of protein translation (**Fig. 1.6d**). For an activation system, reducing the strength of the activator's RBS can reduce the basal expression level of its cognate promoter (Wang et al. 2014; Rubens et al. 2016). For a repression system, regulating the repressor's RBS can change its promoter's input/output dynamic range (Wang et al. 2011a; Nielsen et al. 2016). By changing the RBS of an output of a sensor, its output dynamic range will be directly regulated (Wang et al. 2011a; Wang and Buck 2014). A few libraries of RBS have been summarised and tested in prokaryotes, and can be directly used for tuning translation (Registry of Standard Biological Parts 2017). Moreover, RBS with predictable strength can also be designed easily by using an RBS calculator (Salis et al. 2009).

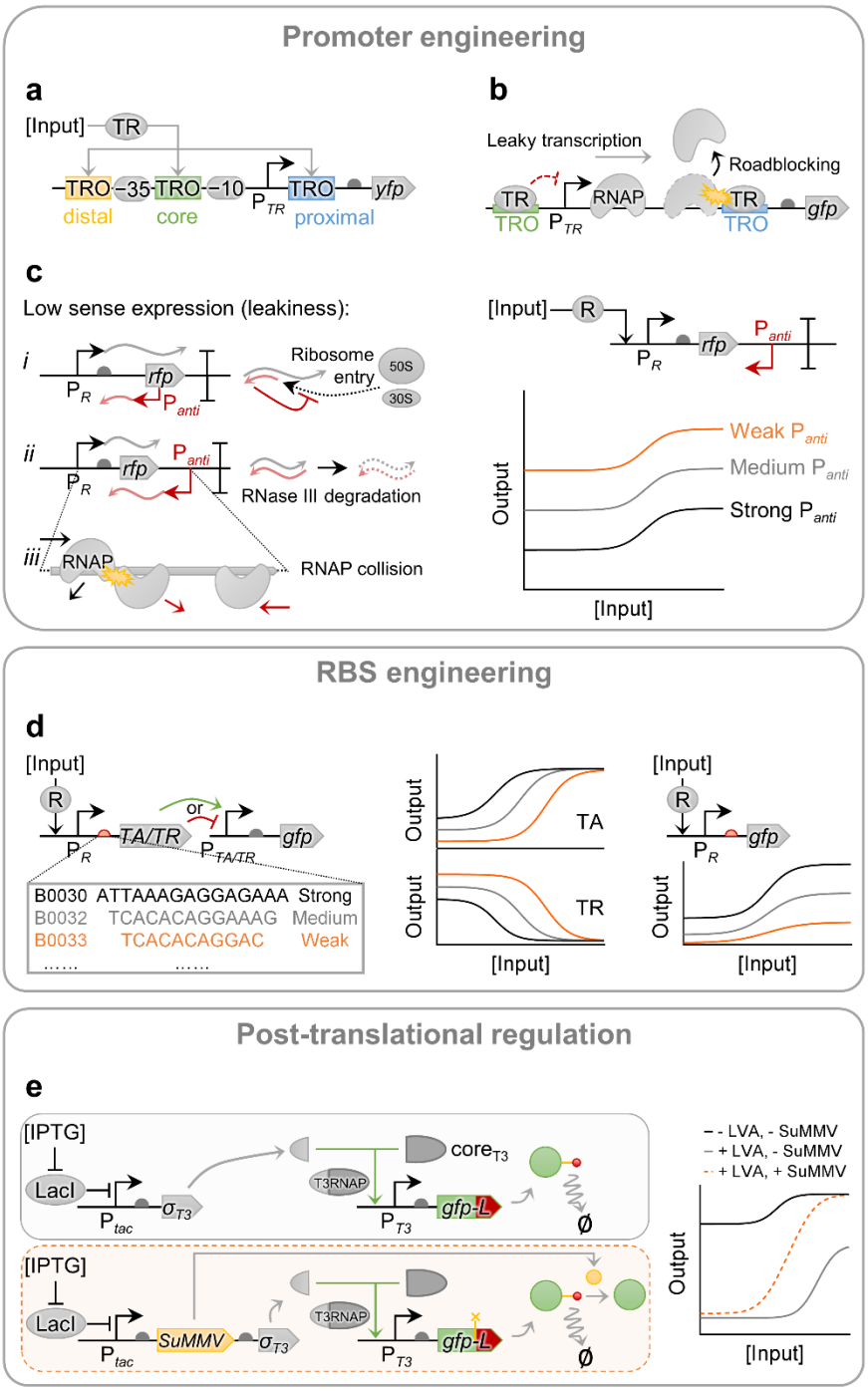


Figure 1.6: Strategies for tuning biosensor leakiness and output dynamics.

(**a**) A transcriptional repressor (TR)-based inducible promoter, with an operator site (TRO) at the distal, core or proximal region of the promoter. Repression efficiency was shown to depend on the TRO's location, with core ≥ proximal ≥ distal (Cox et al. 2007). P_{TR}, TR's cognate promoter. *yfp*, gene encoding a yellow fluorescent protein. (**b**) Transcriptional roadblocking effect. *gfp*, gene encoding a green fluorescent protein. (**c**) Antisense transcription as a tool to tune gene expression. Left panel: antisense promoter (P_{anti}) can reduce P_R's leakiness by blocking the ribosome entry to a reporter's mRNA (**c-*i***),

Apart from RBS sequences, numerous factors have been shown to affect translation rates, such as sequences between the RBS and the gene of interest (Kosuri et al. 2013; Mutalik et al. 2013), additional sequence upstream and downstream of the RBS (Salis et al. 2009; Wu et al. 2018), and codon usage. For example, gene expression can be down-regulated by increasing A/T rich repetitive sequences between the RBS and the gene (Egbert and Klavins 2012), or introducing a long sequence with low GC content between the promoter and the gene (Wu et al. 2018). These factors can be additional targets for regulating the leakiness at translational level.

1.3.5.3 Managing leakiness at post-translational level

Regulating sensor output directly can be a straightforward way to tune sensor leakiness, such as by tuning the RBS of the output (**Fig. 1.6d**) and using an antisense promoter to regulate the output translation (**Fig. 1.6c**). A protein degradation tag can also be used to decrease regulator or output expression level (Andersen et al. 1998; Arpino et al. 2013; Cameron and Collins 2014; Bradley et al. 2016b). However, these methods may shift down the maximum output level, as well as sensor sensitivity (**Fig. 1.6c**–**e**).

An alternative way has been recently proposed to reduce sensor leakiness while maintaining maximum output expression. Fernandez-Rodriguez and Voigt (2016) designed a *Potyvirus* protease-mediated regulation system. In an IPTG sensor without protease regulation, the IPTG inducible promoter P_{tac} solely controlled σ_{T3} expression, which then joined a T3 RNAP core to form an entire T3 RNAP; the entire T3 RNAP activated P_{T3} to produce superfolder GFP. This sensing system showed a high basal expression level and low output dynamic range (**Fig. 1.6e** black line). To reduce the leakiness, an LVA degradation tag was fused to the N terminus of the GFP, which led to less leakiness but also low output expression upon IPTG induction (**Fig. 1.6e** grey line). In order to rescue the output while maintaining the low basal level, a protease (SuMMV) cleavage site was inserted between the GFP and LVA tag, and a protease SuMMV was placed in the same operon as σ_{T3} (**Fig. 1.6e**). By doing this, when induced with IPTG, the SuMMV was expressed and cut off the LVA tag, and the GFP level was fully brought back to

triggering mRNA degradation (**c-*ii***), or blocking the RNA polymerase (RNAP) reading from the sensing promoter P_R (**c-*iii***). Right panel: The sensor's leakiness and output dynamics correlate with the strength of P_{anti} (Brophy and Voigt 2016). R, receptor. P_R, receptor's cognate promoter. (**d**) Modification of ribosome binding site (RBS) for a transcriptional activator (TA), a TR, or the output reporter changes the leakiness and output dynamics of a biosensor (Wang et al. 2011a; Nielsen et al. 2016; Rubens et al. 2016). P_{TA}, TA's cognate promoter. (**e**) Schematics of post-translational regulation on an IPTG biosensor (Fernandez-Rodriguez and Voigt 2016). *L* represents a protein degradation tag, LVA.

the original output level (**Fig. 1.6e** orange dotted line). As such, the IPTG sensor was less leaky but maintained high output expression. This protease-based regulation method is flexible, and can be easily and quickly applied to any other leaky sensing systems.

1.4 Functional expansion of biosensors by synthetic biology

1.4.1 Memory devices

Memorising the occurrence or concentration of a sensing target is sometimes important, especially for clinical diagnoses when samples cannot be directly reached, or when a certain situation need to be determined from the recent past (Courbet et al. 2015). By incorporating a memory circuit into a biosensing circuit, the biosensor will be able to record its target signal and stabilise the output for long term. Synthetic biology communities have developed a number of memory devices (Inniss and Silver 2013; Roquet and Lu 2014; Bradley and Wang 2015). In this section, I will only cover some well-known examples that have been successfully demonstrated in biosensors.

1.4.1.1 Toggle-switch based memory devices

The synthetic toggle-switch can be considered as the earliest reversible memory device. It achieves bistability by using two repressible promoters and their repressors arranged in a mutually inhibitory network (**Fig. 1.7a**) (Gardner et al. 2000). It can flip between two stable states by transient induction which changes the ratio between the two repressors. For example, a biosensor with this toggle switch can be originally in the TR2 state while the TR1 state is repressed (**Fig. 1.7a** Memory off); upon sensing its target, the sensing promoter expresses more TR1 and increases the ratio of TR1 to TR2 (**Fig. 1.7a** Environmental stimulus); higher TR1 will repress the TR2 state and flip the switch to the TR1 state, therefore the sensor will continuously express the output involved in this state (**Fig. 1.7a** Memory on). As only two states are allowed, any graded input response can be converted by toggle-switches to a digital output, so the target concentrations cannot be titrated; however, this may improve sensor robustness and aid decision making processes (Roquet and Lu 2014).

A common toggle-switch is based on a phage lambda cI/Cro genetic switch, and it has been widely used in biosensors due to its high repression efficiency and modularity; but other strong repressors can also be used. The cI/Cro toggle-switch has been used to engineer *E. coli* to sense and record antibiotic exposure or inflammation in murine guts (Kotula et al. 2014; Riglar et al. 2017). A similar toggle-switch has also been used to develop a Pavlovian-like conditioning circuit in *E. coli*, where the toggle-switch helped to memorise a conditioned stimulus (Zhang et al. 2014).

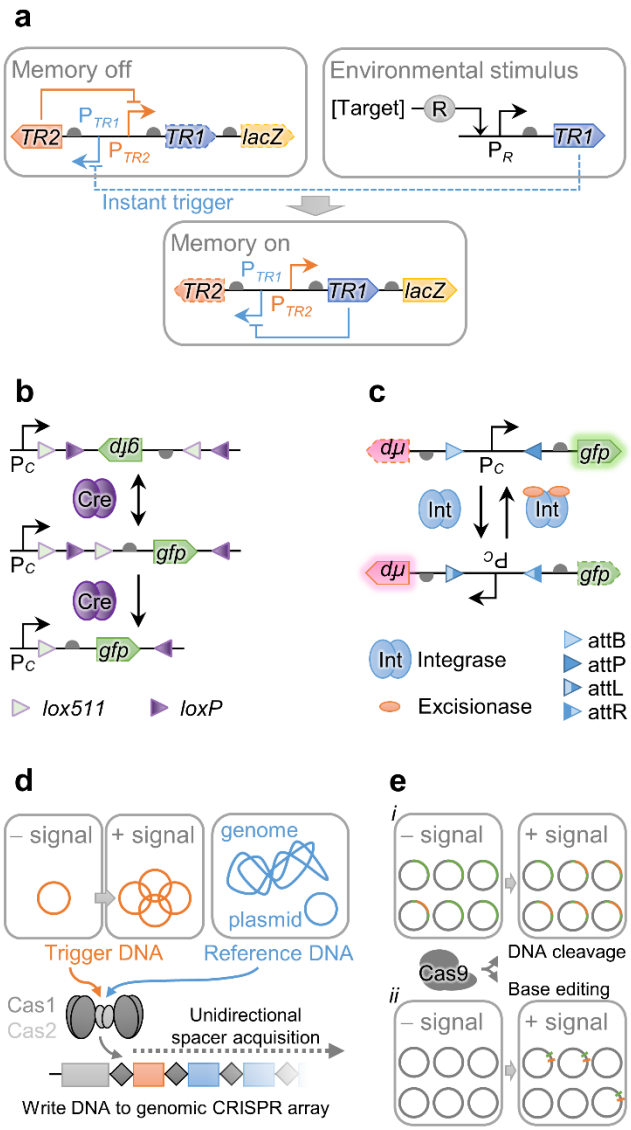


Figure 1.7: Diagrams of genetic circuits for memorising environmental signals.

(**a**) A toggle switch-based memory device. Initially the toggle switch is in the OFF state, where TR2 is expressed and TR1 expression is repressed. Upon target detection, the sensing circuit expresses TR1, which flips the device into the ON state. TR1 and LacZ will be continuously expressed even after the target is removed (Kotula et al. 2014; Riglar et al. 2017). R, receptor. P_R, R's cognate promoter. TR, transcriptional repressor. P_{TR}, TR's cognate promoter. *lacZ*, gene encoding β-Galactosidase for colorimetric output. (**b**) A recombinase-based memory device. Upon sensing a particular target, the biosensor produces recombinase Cre, which first flips the orientation of the *gfp* flanked by *loxP* sites, and then excises one of the two *loxP* sites through the *lox511* sites (Schnütgen et al. 2003). P_C, constitutive promoter. *gfp*, gene encoding a green fluorescent protein. (**c**) An integrase-based memory device switches the sensor's output from *gfp* expression to *rfp* expression. The integrase and excisionase together restore the *gfp* expression (Bonnet et al. 2012). *rfp*, gene encoding a red fluorescent protein. (**d**) A CRISPR-based ''biological tape recorder' system. The signals are recorded into the genomic

1.4.1.2 Recombinase based memory devices

Site specific recombinases are enzymes that perform either DNA excisions or inversions at specific recognition sites (Olorunniji et al. 2016). They belong to either tyrosine or serine recombinase classes, and can be further classified by their directionality and mode of actions (Wang et al. 2011b). For some recombinases that perform both excision and inversion (e.g., Cre), the action depends on the relative orientations of the recognition sites; aligned sites lead to excision while anti-aligned sites lead to inversion (**Fig. 1.7b**). Both allow memory functions if the target DNA is within a functional biological part (e.g., promoter, terminator and gene), but the latter is reversible (**Fig. 1.7b**). The reversible inversion tends to create a mixed population of cells in either ON or OFF states, which reduces the robustness of the memory (Schnütgen et al. 2003; Brophy and Voigt 2014). Solutions include mutagenesis at the recombination sites (Albert et al. 1995; Oberdoerffer 2003), or introducing configurations that leads to the removal of the recombination sites after inversion (**Fig. 1.7b**) (Schnütgen et al. 2003).

Serine integrases are a family of recombinases derived from bacteriophages, which show preferred directionality; therefore the inversion guided by serine integrases is irreversible (**Fig. 1.7c**) (Stark 2017). The irreversible recombinases also have the same mixed population issue when their expression is at intermediate levels (Moon et al. 2011; Brophy and Voigt 2014); however, the issue can be alleviated via the use of feedback loops (Folliard et al. 2017).

Due to the unidirectionality, serine integrases are widely used to generate irreversible memory devices; moreover, the modified genetic context is heritable and can persist even after cell death (Siuti et al. 2013). A simple memory device can be built by inserting a terminator flanked by integrase recognition sites between a promoter and a translational unit, which will interrupt the transcription unless it is excised or flipped by integrases (Bonnet et al. 2013). Other functional parts like promoter and gene can also be designed in this way to generate memory devices. To date, a large number of orthogonal integrases have been identified and tested in cells (Yang et al. 2014). They have been demonstrated to record inflammations in gastrointestinal tracts as well as pathogenic markers in human serum and urine samples

CRISPR array (Sheth et al. 2017). When there is no signal, only the reference DNA will be recorded; where there are signals, the trigger DNA will be rapidly replicated and preferentially recorded into the CRISPR array. (**e**) The CAMERA recording system (Tang and Liu 2018) has two possible mechanisms: ***i***, it uses Cas9 nuclease to record signals by shifting the ratio of two recording plasmids; ***ii***, it uses Cas9-derived base editors to change DNA sequences upon sensing a signal.

(Archer et al. 2012; Courbet et al. 2015; Mimee et al. 2015). Moreover, by coordinating multiple integrases in a memory device, the sensor can record a sequence of detection events and report a unique state to every possible orders (Roquet et al. 2016).

With irreversible memories, sensors would be single use, therefore they may not be suitable for some cases. However, conditional bidirectionality can be achieved by introducing recombinase-excisionase pairs which can make the memory devices as reusable and rewritable as toggle-switches (**Fig. 1.7c**) (Bonnet et al. 2012; Bonnet et al. 2013). Also, similar to toggle-switches, recombinase-based memory devices can be considered as analog to digital converters as the inverted DNA can only be in either of the two states (Rubens et al. 2016).

1.4.1.3 CRISPR/Cas-based memory devices

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) and CRISPR associated system (Cas) are bacterial immunity systems against phage infections (Shalem et al. 2015; Jiang and Doudna 2017). They have recently become another popular way for recording signals in biosensors. One example, called a 'biological tape recorder', is able to record multiple metabolites over time (**Fig. 1.7d**) (Sheth et al. 2017). This system is based on a naturally occurring bacterial CRISPR-Cas adaptation process, where the bacteria capture exogenous DNA from invading plasmids/phages and integrate them into their genomic CRISPR arrays as spacers. In the 'biological tape recorder' system, a new spacer is incorporated into the CRISPR array upon the induction of a specific target, and multiple events can be recorded over time.

Another example, named CAMERA (CRISPR-mediated Analog Multi-event Recording Apparatus), either uses Cas9 nuclease to record signals by shifting the ratio of two recording plasmids (based on the DNA cleavage function of Cas9 nuclease to remove one of the plasmids upon sensing), or uses Cas9-derived base editors to change DNA sequences upon sensing a signal (**Fig. 1.7e**) (Tang and Liu 2018). In both examples, the temporal and lineage information is analysed by sequencing the barcoded DNA sequences, or by coupling the resulting change to an observable output.

1.4.2 Signal integration

Generally, biosensors are built to sense a single input, but in many cases the biosensors are required to report a specific condition which may involve multiple inputs, such as some health condition defined by multiple biomarkers. Synthetic biology has provided numerous tools for bio-computation, which can be applied to biosensors to receive and process multiple inputs, therefore to sense and react to a complex condition or global environment (Wang et al. 2011a;

Moon et al. 2012; Wang et al. 2013; Wang and Buck 2014; Ma et al. 2016; Nielsen et al. 2016; Roquet et al. 2016; Rubens et al. 2016).

Up to now, although no genetic circuits have been developed in bacterial sensors to target multiple clinical biomarkers, several studies have been done to demonstrate the possibilities, such as using AND logic gates to sense three different metal ions (**Fig. 1.8**) (Wang et al. 2013), and engineering *E. coli* to invade tumour cells only when they detect hypoxia and reach high cell density (Anderson et al. 2006).

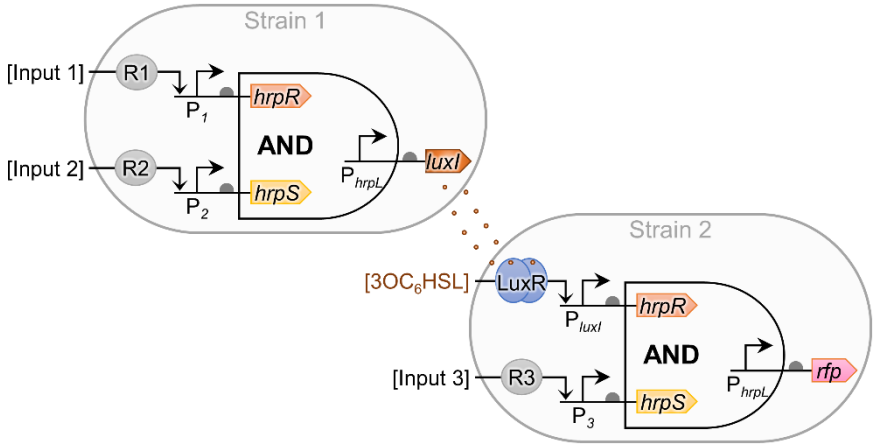


Figure 1.8: Detecting an environmental condition using multi-input AND gate and cell-cell communication.

A three-input AND logic gate based on two HrpRS-based AND gates separated in two different strains in a consortium (Wang et al. 2013). *luxI*, synthase of a quorum sensing molecule $3OC_6HSL$. *rfp*, gene encoding a red fluorescent protein.

1.4.3 Reshaping the response function

Most biosensor response curves are sigmoidal, which is either analog or digital; for the former, the output relies on graded input, so the target concentration is titratable, while the latter has clear ON and OFF states therefore is robust for decision making purposes. The analog response can be easily converted to a digital response by using the aforementioned integrases and some repressor-based amplifiers (Hooshangi et al. 2005; Rubens et al. 2016), and a semi-log response curve can be generated by using a coherent feedforward loop (Zhang et al. 2013). Further, incoherent feedforward loops based on transcriptional regulators or integrases have been developed to build bandpass filters in biosensors, which can only response to a limited range of target concentrations (Peking iGEM 2013; Rubens et al. 2016).

1.5 From the lab to the field

Combining synthetic biology and cellular sensors have shown their promising abilities in a broad range of applications, such as environmental diagnosis and bioremediation (Trang et al. 2005; Buffi et al. 2011; De Mora et al. 2011; Siegfried et al. 2012; Tong et al. 2014; Huang et al. 2015a; Rasmussen and Minteer 2015; Wu et al. 2017), clinical diagnosis and biotherapy (Saeidi et al. 2011; Hwang et al. 2014; Courbet et al. 2015; Danino et al. 2015; Riglar et al. 2017; Watstein and Styczynski 2018), bioprocessing (Zhang and Keasling 2011; Zhang et al. 2012), mining and landmine detection (Cerminati et al. 2011; Belkin et al. 2017). However, using engineered organisms beyond controlled laboratory conditions has raised safety concerns, which limit their applications.

The potential risk of using engineered organisms is to disrupt the balance of nature, which can be classified in four areas: **1)** they may interact with and be harmful to the surrounding environment; **2)** if they escape from their pre-targeted places, they may survive and change natural habitats, food webs or biodiversity; **3)** they may evolve and adapt to fill new ecological niches; and **4)** they may transfer their genes to native organisms, especially antibiotic resistance genes which may cause threats to human health (Dana et al. 2012). To avoid natural disasters in the future, these biosafety issues should always be considered before designing and applying the new generation of engineered organisms. To address these limitations, both intrinsic containment mechanisms and physical containment systems have been studied and tested for synthetic organisms (Moe-Behrens et al. 2013; Wright et al. 2013; Wright et al. 2015).

1.5.1 Intrinsic containment

1.5.1.1 GeneGuard

Wright and his colleagues developed a stable and modular system ‘GeneGuard’ for biosafety control in *E. coli* (Wright et al. 2015). Three safety modules were combined in this system, including host-dependent origins of replication, rich-media compatible auxotrophy, and toxin-antitoxin pairs. The first module consists of conditional origins of replication, which means the plasmid replication is only possible in engineered host cells. The second module is based on auxotrophy as a selective marker. The engineered cells can only survive with specific supplements which will not exist in nature, so the engineered cells cannot survive if they escape from the controlled environment (e.g., industrial closed systems). The last module carries of designed toxin-antitoxin pairs, which work in an auxiliary manner to prevent gene propagation.

1.5.1.2 Linguistic barrier

A 'linguistic barrier' is an attractive way to eliminate biosafety issues from engineered cells (Wright et al. 2013). It involves refactoring transcriptional or translational machineries, making them incompatible with wild microorganisms. One approach is to refactor the translation machinery in engineered hosts, so that only the hosts can correctly translate the synthetic sequences. To achieve this, the codon assignments can be refactored in the cells. Another way is to evolve ribosomes which can recognise non-natural RBS or a quadruple-base-pair code. As such, even when the synthetic genes are acquired by wild cells, they cannot use their own ribosomes to translate the genes properly. Other methods have also been proposed to disrupt the transcriptional language, such as using alternative base pair combinations, and creating RNAP mutants to recognise xeno nucleic acids (Pinheiro et al. 2012).

1.5.1.3 Death induction

The next generation of biosafety devices may also involve coupling synthetic counting circuits (Friedland et al. 2009; Callura et al. 2010) or a predictable timer (Ellis et al. 2009) to induce cell death via expression of toxin components or other cell death pathways. As such, the engineered cells would commit suicide after a pre-defined retention time in the environment. Moreover, cell death induced by environmental signals is also possible by combining a signal-inducible promoter and suicide genes, such as for engineered cells escaping from an anoxic environment (e.g., tumour) to an aerobic environment (e.g., normal tissue), or from high temperature (e.g., specific culture conditions) to low temperature (e.g., normal environment). However, those methods still have evolutionary risks, such as that inactivation of the suicide mechanism by mutation would allow the cell to escape.

1.5.2 Physical containment

Although several studies have been carried out to prevent potential genetic pollution from engineered organisms, the issue still remains of how to prevent the synthetic cells from competing with native cells and destroying the ecological balance. Unfortunately, even if scientists find genetic solutions to address these issues, it would be hard to test their effectiveness in synthetic cells as some problems might only occur after prolonged evolution. Thus, physical containment is always indispensable for biosafety control, particularly for cellular biosensors in field testing and bioprocessing.

Several kinds of physical containment have been developed to limit the movement of synthetic organisms, such as confining the cells to bioreactors for biosynthesis applications in factories,

or using microencapsulation to limit the mobility of the cells, e.g., in alginate beads, agarose and silica gels (Chang and Prakash 2001; Nassif et al. 2002; Papi et al. 2005; Sharma et al. 2010; Buffi et al. 2011; Power et al. 2011; Shin 2012; Courbet et al. 2015; Belkin et al. 2017), and in bioreactive nanofibers (Tong et al. 2014). It has been shown that by entrapping the cells in hydrogels, the synthetic cells can be functional for at least a month, and cells can be viable after one year's storage. However, better entrapping materials and storage conditions still need to be discovered to maintain sensor viability and functionality for longer periods.

Recently, a microfluidic device has been developed to entrap cellular sensors for field testing (Volpetti et al. 2017). This device contains 768 biopixels, and each biopixel contains a chamber where sensor cells can be spotted. Each chamber also has valves around it to control the flow of media or samples to the sensors, and also to entrap the sensors in the chamber. If the sensors have fluorescent output, images can be taken to show the sensor response. This device has been used to test arsenic and arabinose inducible sensors, and images showing sensor response can be taken using a microscope or portable device like a USB microscope or a mobile phone. To enable long-term storage of cellular sensors in this microfluidic device, spores of *B. subtilis* containing an IPTG sensing circuit have been stored and tested; after incubating them at 80°C for a month, the sensors still can show their functionality and survivability. However, using the microfluidic device still requires a medium pumping and flow system, which may increase the practical difficulties for screening of large numbers of samples.

1.5.3 Cell-free biosensors

1.5.3.1 In vitro transcription-translation

Cell-free systems (CFS) have been proposed as a solution to circumvent the biosafety issues associated with whole-cell biosensors. Cell-free biosensors are basically sensory genetic circuits in a CFS.

The CFS I mention here is an *in vitro* transcription-translation system, where there are no cells but only biological machinery and energy source required for transcription and translation. This system can be either based on crude cell extract (Kim et al. 1996; Kigawa et al. 2004; Kwon and Jewett 2015; Didovyk et al. 2017) or a system of purified recombinant elements (PURE) necessary for transcription-translation (Shimizu et al. 2001; Kuruma and Ueda 2015). The former is cheap and easy to produce, but contains unknown background components from the cell genome. The latter contains well-known components, but is more expensive and time consuming to produce. Both have been widely used for different purposes, such as rapid protein expression, rapid prototyping of genetic circuits, rapid characterisation of molecular

interactions and biosensing (Pellinen et al. 2004; Chappell et al. 2013; Caschera and Noireaux 2014; Hong et al. 2014; Li et al. 2014; Pardee et al. 2014; Siegal-Gaskins et al. 2014; Sun et al. 2014; Niederholtmeyer et al. 2015; Takahashi et al. 2015a; Takahashi et al. 2015b; Wu et al. 2015; De Los Santos et al. 2016; Garamella et al. 2016; Karim and Jewett 2016; Pardee et al. 2016a; Duyen et al. 2017; Wen et al. 2017; Halleran and Murray 2018; Marshall et al. 2018; Maxwell et al. 2018; Sansbury et al. 2018). Most studies were carried out in the cell extract-based CFS (CE-CFS), which may be because it is cheap and can be produced easily in the laboratory. Moreover, the process of CE-CFS production has been recently optimised and significantly simplified, which makes it much easier and more affordable (Kwon and Jewett 2015; Didovyk et al. 2017). Notably, compared to PURE, CE-CFS is more similar to the intracellular environment, which makes it a more predictable way to transfer designed genetic circuits between whole cells and CE-CFS.

Particularly for biosensing, CFS-based sensors can be more tolerant to toxic translational products than cells, more sensitive and the response to their target can be faster. This is because the CFS does not contain a physical barrier (e.g., cell wall/membrane and transport systems) to dilute the targets and slow down the targets reaching the sensing elements. Also, it does not need to maintain cell growth like a whole cell, so toxic proteins that can cause cell burden will have less or no effect in CFS. Due to these advantages, more and more CFS-based sensors have been developed recently to detect heavy metals, quorum sensing molecules from pathogens, antibiotics and viral RNA (Pellinen et al. 2004; Pardee et al. 2014; Pardee et al. 2016a; Didovyk et al. 2017; Wen et al. 2017).

Current cell-free reactions are mostly carried out in solution so that test tubes are required, which are inconvenient for large scale field testing. However, Pardee et al. (2014) recently developed a paper-based sensing platform, which can be freeze-dried and stay stable over a year. This platform basically adapts a CFS reaction on a piece of paper instead of using a test tube. Notably, this platform is portable and can be easily handled, which will be convenient for field testing (Pardee et al. 2016b). More importantly, the paper-based sensing system is cost effective as only 1.8 μl cell-free reagent is needed for each reaction; also, by using homemade cell extract, it is possible to reduce the cost below $1 for each test. Both fluorescent and colorimetric output can be used in the paper-based sensors. Although Pardee et al. (2014) have developed a cheap electronic reader ($ 250) for quantifying sensor output, it will be possible to just use a smart phone to quickly image and quantify sensor output, which will make field testing even more convenient and more affordable (Wei et al. 2014).

Although the CFS-based sensors have those practical advantages, especially for paper-based sensors, there are still some drawbacks which need to be considered. First, the genetic circuits may not work in the same way in a whole cell and in the CFS, which would increase the difficulties in transferring the sensors from *in vivo* to *in vitro*. This may be caused by two effects: **1)** CFS does not have a physical barrier, which makes it lose the macromolecular crowding properties, and therefore changes protein conformation, binding of small molecules, protein-protein and protein-nucleic acid interactions, etc. (Shahid et al. 2017); **2)** even for the CE-CFS, its components will not be exactly the same as those in a whole cell, therefore the resource allocated to the sensory circuits will be different from that in the cells. The macromolecular crowding can be potentially achieved by using minimal/artificial cells (i.e., CFS entrapped in liposomes), and the component variance can be minimised by optimising the CFS preparation (Shin and Noireaux 2012; Tan et al. 2013; Garamella et al. 2016); however, further tuning will still be required for different cases. Secondly, CFS may be not as stable as whole cells. This is because the CFS need to be reassembled each time, and small variance at each step may affect the final activities of the CFS. Therefore, careful preparation is required for each batch of CFS, and an activity calibration will be essential (Takahashi et al. 2015b). Last, testing conditions for CFS may still be a practical limitation, especially for the paper-based sensing system which would need to work in a humidified environment to prevent evaporation.

1.6 Subjects and objectives

Although many synthetic biology strategies have been studied individually to improve cell-based biosensors, each of them still has its own limitations. For example, simply reducing the ArsR intracellular density also increased the leakiness of the arsenic sensor (Wang et al. 2015), and the HrpRS-based amplifier could not significantly improve the output of the arsenic sensor when the arsenic level was $\leq$ 10 ppb (Wang et al. 2014). More importantly, very few cell-based biosensors have undergone field trials, suggesting that further optimisation is still needed for most of these sensors in terms of sensing performance and practicability. To date, however, no systematic method has been created to guide the improvement of cell-based biosensors, and there has been no cheap and easy-to-use platform developed for sensor storage and large sample screening.

My PhD project aims to develop a novel, comprehensive and modular methodology for optimising cell-based biosensors to overcome their practical difficulties, and also to prepare them for their projected applications. In particular, this methodology will combine multiple synthetic biology strategies, which can systematically and significantly improve sensing

performance in a predictable manner. It will first optimise sensor sensitivity by regulating intracellular receptor densities, then further improve sensor output by applying a novel multi-layer amplifier cascade, and finally regulate sensor leakiness by combining promoter structure engineering and post-translational regulation. Previously studied arsenic and mercury sensors will both be used to test the practicality and modularity of this methodology.

To prepare the arsenic and mercury sensors for real-world applications, I will also develop a sensing platform (e.g., a sensor array) for inexpensive, portable and easy-to-use field testing where analytes can be easily quantified by simply visualising the output patterns from the sensor array. Physical entrapment methods will be applied to address the biosafety concerns. To further minimise the biosafety issues, similar sensing platforms will be adapted in a paper-based CFS.

Chapter 2. Materials and Methods

2.1 Materials

2.1.1 Growth media

All the media used in this study were sterilised by autoclaving at 121°C for 15 min or by filtration using 0.22 μm syringe filters (SLGP033RS, Millipore). For M9 minimal medium, all the components were sterilised separately and mixed prior to use. Detailed ingredients, application and sterilisation methods of different media are summarised in **Table 2.1**.

Table 2.1: Bacterial growth media used in this study

Medium	Ingredient	Used for	Sterilisation method
Lysogeny broth (LB)	10 g/L peptone, 5 g/L NaCl, 5 g/L yeast extract	General bacterial culture in liquid media	Autoclaving
LB agar	10 g/L peptone, 5 g/L NaCl, 5 g/L yeast extract, 16 g/L agar	General bacterial growth on solid media	
M9 minimal medium with 0.4% glycerol	11.28 g/L M9 salts (M6030, Sigma-Aldrich), 1 mM thiamine hydrochloride, 0.2% (w/v) casamino acids, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.4% (v/v) glycerol	Incubation of agarose-entrapped cells	Filter-sterilisation for thiamine, autoclaving for the others
2 × M9 minimal medium with 0.8% glycerol	22.56 g/L M9 salts, 2 mM thiamine hydrochloride, 0.4% (w/v) casamino acids, 4 mM $MgSO_4$, 0.2 mM $CaCl_2$, 0.8% (v/v) glycerol	Incubation of agarose-entrapped cells	
M9 minimal medium with 0.4% glucose	11.28 g/L M9 salts, 1 mM thiamine hydrochloride, 0.2% (w/v) casamino acids, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.4% (w/v) glucose	Incubation of agarose-entrapped cells	

2.1.2 Antibiotics

Antibiotics used for bacterial growth are listed in **Table 2.2**. Apart from chloramphenicol, all of the antibiotics were sterilised by filtration using 0.22 μm syringe filters.

Table 2.2: List of antibiotics

Antibiotic	Solvent	Stock concentration	Working concentration
Ampicillin	ddH_2O	100 mg/ml	100 μl/ml*
Kanamycin	ddH_2O	50 mg/ml	50 μl/ml
Chloramphenicol	Ethanol	25 mg/ml	25 μl/ml

*: Ampicillin working concentration was 50 μl/ml for cells harbouring low copy number plasmid.

2.1.3 Bacterial strains

Two *E. coli* strains were used in this study and details are listed in **Table 2.3**.

Table 2.3: *E. coli* strains used in this study

Strain	Genotype	Used for	Source
TOP10	F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lacZ*ΔM15 Δ*lacX*74 *nup*G *recA*1 *araD*139 Δ(*ara-leu*)7697 *gal*E15 *gal*K16 *rps*L(Str^R) *end*A1 λ^-	General cloning and gene expression assays	Laboratory collection
MG1655	K-12 F^- λ^- *ilvG*$^-$ *rfb-50 rph-1*	Genomic DNA as PCR template	

2.2 Strains, plasmids and growth conditions

2.2.1 Strains and growth conditions

Plasmid cloning and genetic circuit characterisation were all performed in *E. coli* TOP10. Cells were cultured in LB medium, with appropriate antibiotics. For circuit characterisation, the engineered bacteria were first inoculated from a single colony on a freshly streaked solid LB plate to 2 ml LB medium, and cultured overnight at 37°C with shaking (200 rpm). Then the cells were diluted 100-fold from the overnight culture into fresh LB medium. For liquid culture induction, the diluted culture was loaded into a black 96-well microplate with clear bottom (655096, Greiner Bio-One) and induced with 5 μl inducers to a final volume of 200 μl per well. For testing the constitutive promoters, 200 μl of diluted culture was loaded to each well. The microplate was sealed with an air permeable film (AXY2006, SLS), and incubated in a shaker incubator (AS-03020-00, Medical Supply Co. Ltd) with continuous shaking (1,000 rpm, 37°C). After 5 h unless otherwise indicated, a fluorescence plate reader (BMG FLUOstar) was used to monitor cell growth and measure fluorescence.

Antibiotics and other reagents used for induction assays (i.e., $HgCl_2$, $NaAsO_2$, Na_2HAsO_4, $CuSO_4$, $MgSO_4$, $ZnCl_2$, $FeCl_2$, $NiCl_2$, $MnCl_2$, $PbCl_2$, $CoCl_2$, and $CdCl_2$) were analytical grade and purchased from Sigma-Aldrich. The reagents in solid form were dissolved in ddH2O and were then filtered using 0.22 μm syringe filters (SLGP033RS, Millipore).

2.2.2 Plasmid construction

Standard molecular biology techniques were used to construct plasmids containing arsenic or mercury-responsive genetic circuits and amplifiers. Sequences of genetic parts used are listed in **Appendix Table 2.1**, plasmids used in this study are summarised in **Appendix Table 2.2**, and detailed plasmid maps are provided in related figures and **Appendix Figure 3.1**. All primers used in this study are listed in **Appendix Table 2.3**.

For tuning the intracellular receptor densities, the constitutive promoters and the ribosome binding sites (RBS) which directly drive the expression of *arsR*/*merR* were replaced via Polymerase Chain Reaction (PCR) amplification (see Section 2.3.2). To optimise the amplifier $Amp30^C$, the σ^{54} binding site in P_{hrpL} was replaced with a σ^{54} consensus binding sequence (**Fig. 3.3**) by PCR.

The *rinA_p80α* and *ECF* genes, and promoter $P_{rinA\_p80\alpha}$ were synthesised by GeneArt following the BioBrick™ standard (http://biobricks.org) (see Section 2.3.1). The genes were separately constructed with a terminator (BBa_B0015) via BioBrick assembly (see Section 2.3.9) (Shetty et al. 2008). Various RBS sequences and a degradation tag sequence (e.g., ASV tag) were introduced to the gene by PCR. Cognate promoters for ECFs were constructed by annealing oligonucleotides.

P_{arsR} variants with an extra ArsR binding site (ABS) at varying distances or with different σ^{70} binding sites were constructed by inverse PCR (see Section 2.3.2) (Ochman et al. 1988). Genes *malE* and *TEV* were amplified from the *E. coli* MG1655 genome and BBa_K1319004 separately by PCR, and were assembled together by Gibson Assembly (see Section 2.3.10) (Gibson et al. 2009). The fused sequence was next ligated with P_{arsR}-ABS62 into pSB4A3 by BioBrick Assembly. The TEV protease cleavage sequence was added between *gfp* and *AAV* tag sequence by inverse PCR. An extra linker sequence was then added ahead of the cleavage site by a second round of inverse PCR.

2.3 DNA manipulation, purification and analysis

2.3.1 DNA synthesis and sequencing

DNA sequences to be synthesised by GeneArt were designed following the BioBrick standard: the DNA was designed to be flanked with BioBrick prefix and suffix, and to remove the

BioBrick standard restriction sites (i.e., EcoRI, XbaI, SpeI and PstI) if they were within the sequences of the gene/promoter.

Oligonucleotides for PCR (see Section 2.3.2) and annealing (see Section 2.3.7) were synthesised by Sigma or IDT. Oligonucleotides longer than 50 bp or for inverse PCR were ordered from Sigma with cartridge purification.

The plasmids constructed in this study were sequenced by Edinburgh Genomics, MRC PPU DNA Sequencing and Services or Source BioScience. The DNA samples were prepared following the services' instructions. The sequencing results were analysed by FinchTV or SnapGene software. Oligonucleotides for sequencing are listed in **Appendix Table 2.3**.

2.3.2 Polymerase chain reaction (PCR)

Applied Biosystems Verti™ Thermal Cycler or ProFlex PCR System was used to perform PCR. DNA templates were prepared from purified DNA (diluting to 1 ng/μl, see Section 2.3.5) or cell colonies (single colony was resuspended in 100 μl sterilised ddH_2O). Typically, 1 μl of diluted DNA or cells was used for each 50 μl PCR reaction mixture.

2.3.2.1 General Phusion PCR

To amplify DNA fragments for subsequent cloning, Phusion high-fidelity DNA polymerase (M0530, NEB) was used. PCR was prepared and performed following the manufacturer's protocol. The annealing temperature for each primer pair was estimated using an online Tm calculator (https://www.thermofisher.com). However, for some cases, additives (i.e., DMSO or 16% (v/v) glycerol) were added, and touchdown PCR (Don et al. 1991) or gradient PCR (Prezioso and Jahns 2000) was performed.

2.3.2.2 Overlap extension PCR

Overlap extension PCR was used to joint two DNA fragments for later cloning. It was performed using Phusion high-fidelity DNA polymerase following a previous protocol with some changes (Nelson and Fitch 2011). 0.012 pmol of each purified fragment (with 28 bp overlapping region) were mixed for touchdown PCR (10 cycles, 60°C to 50°C with 1°C reduction in each cycle for the annealing step). Then the reaction was mixed with a forward and a reverse primer to amplify the entire joint fragment for 30 cycles with 50°C for the annealing step, followed by a final elongation step at 72°C for 7 min. Details of other PCR steps followed the manufacturer's protocol. The PCR product was then analysed by agarose gel electrophoresis (see Section 2.3.4). PCR product with correct size was purified (see Section 2.3.5) for restriction digestion (see Section 2.3.3) and ligation (see Section 2.3.7).

2.3.2.3 PCR for site-directed mutagenesis

Site-directed mutagenesis was used to generate small mutations in genes and promoters, such as for optimising P_{hrpL} and engineering arsenic inducible promoters, for inserting protein degradation tag and linkers, and for removing restriction sites from genes of interest. The PCR procedure was the same as the general Phusion PCR.

For single point-mutations, or two point mutations in adjacent regions, the protocol from Zheng et al. (2004) was followed with some changes. A pair of primers were designed with about 16 bp complementary regions covering the mutations (with melting temperature ≤ 72°C) while both primers were complementary to the DNA template with about 20 bp homology. The purified PCR product was directly used for transformation (see Section 2.4.2) after DpnI digestion (see Section 2.3.3).

For insertion or multiple point mutations, these were split as insertions into the paired primers. The blunt-ended DNA products were purified and ligated as described in Section 2.3.5 and 2.3.7.

2.3.2.4 Colony PCR

To identify a bacterial clone with assembled DNA fragments of the correct size, colony PCR (Bergkessel and Guthrie 2013) was performed using GoTaq G2 DNA Polymerase (M7841, Promega) following the manufacturer's protocol. Cells were prepared as described in Section 2.3.2.

2.3.3 DNA restriction digestion

Purified DNA (≤ 1 μg, see Section 2.3.5) was digested in a 10 μl or 20 μl reaction volume with FastDigest enzymes (Thermo Scientific) supplemented with FastDigest Green Buffer. The reaction was prepared following the manufacturer's protocol, and was then incubated at 37°C for 30 min per enzyme. If the DNA fragment was to be used as a vector for later ligation, FastAP Thermosensitive Alkaline Phosphatase (EF0651, Thermo Scientific) was included in the reaction to dephosphorylates DNA ends. The digestion was then analysed by agarose gel electrophoresis (see Section 2.3.4). Details of DpnI digestion prior to transformation are described in Section 2.3.7.

2.3.4 Agarose gel electrophoresis

DNA fragments were separated and analysed by agarose gel electrophoresis. The agarose gel was prepared according to the size of DNA fragments to be separated (e.g., 0.8% (w/v) agarose gel for DNA fragments > 2 kb and 1.6% for DNA fragments < 200 bp). Typically, agarose

(BIO-41025, Bioline) was dissolved in 1 × TBE buffer (20-6000-50, Thistle Scientific) by heating the mixture in a microwave. Afterwards, the melted solution was cooled down to 50 – 60˚C and SYBR Safe DNA Gel Stain (S33102, Life Techonologies) was then added with a dilution of 1 : 10,000. The solution was poured into a gel tank with a comb and was left to solidify. DNA samples from PCR reactions were mixed with 5 × loading dye (BIO-37045, Bioline) before being loaded into the gel while those from restriction digestion were directly loaded into the gel. HyperLadder™ 1kb (BIO-33025, Bioline) was loaded as a molecular weight marker. The gel was run at 125 volts in 1 × TBE buffer for 20 – 40 min until the DNA fragments were well separated. Thereafter, the DNA bands in the gel were visualised using the Bio-Rad Gel Doc XR+ system (with filter 1 and SYBR safe mode).

2.3.5 DNA purification

Plasmid DNA from cell cultures was purified using the Qiaspin Miniprep Kit (27106, Qiagen) following the manufacturer's protocol for general cloning and sequencing. Linear DNA from restriction digestion and PCR was purified using the QIAquick Gel Extraction Kit (28704, Qiagen) or Wizard SV Gel and PCR Clean-Up System (A9282, Promega) following the manufacturers' protocols for subsequent cloning. The DNA was eluted with sterilised ddH_2O or nuclease free H_2O provided by the kits. For cell-free reactions, plasmids were purified using the QIAfilter Plasmid Midi Kit (12243, Qiagen) or ZymoPURE Midiprep Kit (D4201, Cambridge BioScience) following the manufacturers' protocols. The plasmid was dissolved or eluted in nuclease free water (W3500, Sigma-Aldrich). Purified DNA was quantified using NanoDrop 2000 or DeNovix DS-11 spectrophotometer. The size and the quantity of the DNA were also verified by restriction digestion (see Section 2.3.3) and agarose gel electrophoresis (see Section 2.3.4). If necessary, the DNA was concentrated by air drying at 37˚C or by a vacuum concentrator at 45˚C (Concentrator plus, Eppendorf).

2.3.6 Ethanol precipitation

If necessary, ethanol precipitation was used to further purify and concentrate DNA for Gibson assembly and cell-free reactions. The DNA sample (dissolved in water in a 1.5 ml tube) was mixed with 1/10 volume of 3 M potassium acetate (final concentration of 0.3 M), then 2.5 volumes of pre-cooled 100% ethanol (final concentration of 71.4%) was added. The mixed sample was placed at – 20˚C for ≥ 20 min and was then centrifuged at 17,000 g for 15 min. After carefully decanting the supernatant, the sample was washed with 1 ml of pre-cooled 70% ethanol. Afterwards, the sample was centrifuged at 17,000 g for 5 min. After carefully

decanting the supernatant, the sample was briefly air dried or vacuum dried. The DNA pellet was then resuspended in sterilised ddH_2O or nuclease free water.

2.3.7 DNA ligation

Ligations were performed using T4 DNA ligase and 10 X T4 DNA Ligase Reaction Buffer (M0202, NEB) following the manufacturer's protocol. An approximate molar ratio of vector to insert of 1 : 3 was used in each reaction, and the final volume was brought up to 10 µl or 20 µl with sterilised ddH_2O. Reactions were incubated overnight at 4˚C or for 30 min – 1 h at room temperature before transformation (see Section 2.4.2).

For ligating DNA products from blunt-end ligation, 8 µl purified PCR product was incubated with 1 µl T4 Polynucleotide Kinase (M0201, NEB) and 1 µl 10 X T4 DNA Ligase Reaction Buffer at 37˚C for 30 min. Then the reaction was mixed with 8 µl H_2O, 1 µl T4 DNA Ligase and 1 µl 10 X T4 DNA Ligase Reaction Buffer and was incubated for 30 min – 1 h at room temperature. Thereafter, the reaction was mixed with 1 µl DpnI and 2.33 µl 10 X FastDigest Green Buffer at 37˚C for 30 min, followed by inactivation at 80˚C for 5 min. Then the reaction was directly used for transformation or was stored at – 20˚C for later transformation.

2.3.8 Oligonucleotides annealing

Specifically paired oligonucleotides were used to build short double stranded DNA sequences such as promoters (e.g., P_{ecf} and P_{merT}). The oligonucleotides were first phosphorylated separately in a 20 µl reaction (with 10 µM oligonucleotides, 1 X T4 DNA Ligase Reaction Buffer and 1 µl T4 Polynucleotide Kinase), which was incubated at 37°C for 1 h followed by inactivation at 65°C for 20 min. 5 µl each of phosphorylated forward and reverse oligonucleotides were mixed with 90 µl sterilised ddH_2O and were then incubated at 95°C for 3 min. Thereafter, the reaction mixture was cooled down to room temperature for 30 min – 1 h. The annealed product was used for ligation or was stored at – 20˚C for later use.

2.3.9 BioBrick Assembly

Most of the plasmids constructed in this study were generated by the BioBrick standard assembly method (http://parts.igem.org/Assembly:Standard_assembly). A BioBrick part is a DNA fragment of interest designed to be flanked with BioBrick Prefix and Suffix which are comprised of restriction sites: BioBrick Prefix-EcoRI and XbaI restriction sites, BioBrick Suffix-SpeI and PstI restriction sites (**Fig. 2.1a**). BioBrick vectors were also designed including the BioBrick Prefix and Suffix (Shetty et al. 2008), into which BioBrick parts can be easily inserted through restriction digestion and ligation. Further, the XbaI generated sticky

end is complementary to the sticky end generated by SpeI; therefore, two BioBrick parts can be ligated in a BioBrick vector through two way or three way ligation (**Fig. 2.1b,c**).

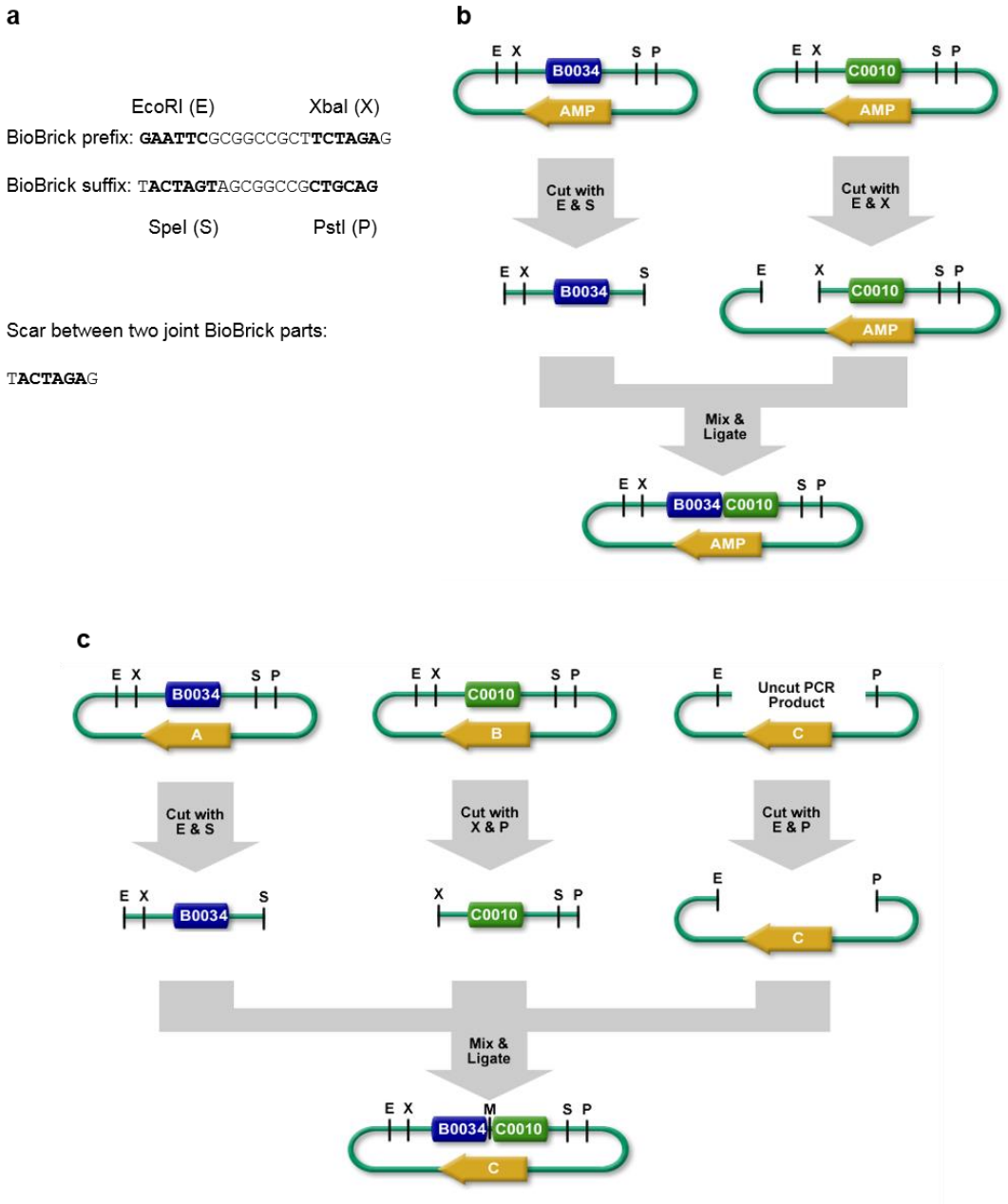


Figure 2.1: Diagrams showing BioBrick assembly.

(**a**) Sequences of BioBrick prefix, suffix and scar. (**b**) BioBrick assembly of two BioBrick parts (http://parts.igem.org/Assembly:Standard_assembly). (**c**) 3-way assembly of three BioBrick parts (http://parts.igem.org/Help:Assembly/3A_Assembly).

This method is easy to use as only four restriction enzymes are involved, and any BioBrick parts can be assembled in a plug-and-play manner; moreover, as the restriction sites are

standardised, the adjacent sequences to the DNA of interest are always the same which can ease the comparison of expression between different constructs. However, BioBrick Assembly always generates an 8 bp scar sequence (**Fig. 2.1a**, ligated from SpeI and XbaI sticky ends) between joined parts, which does not allow generation of fusion proteins.

2.3.10 Gibson assembly

DNA fragments (excluding vector) for Gibson assembly (Gibson et al. 2009) were amplified by PCR using Phusion high-fidelity DNA polymerase, containing overlapping regions of 27 bp or 28 bp. The vector was linearised by restriction digestion and was also dephosphorylated. 0.06 pmol of DNA fragments were mixed with 10 μl NEBuilder HiFi DNA Assembly Master Mix (E2621, NEB) and the final volume was brought up to 20 μl with sterilised ddH_2O. The reaction was incubated at 50˚C for 15 min, and 10 μl was used for transformation. A negative control (comprised of vector DNA and ddH_2O) was performed following the same process.

2.4 Microbiological techniques

2.4.1 Competent cell preparation

Chemically competent cells were prepared and used for heat shock transformation. Reagents (50 mM $CaCl_2$ and 50 mM $CaCl_2$+15% (v/v) glycerol) were autoclaved and stored at 4˚C. A 500 ml Erlenmeyer flask with 200 ml LB was autoclaved and kept at room temperature. All the sterilised tubes, tips and stripettes were pre-cooled to 4˚C or – 20˚C. *E. coli* TOP10 cells were re-streaked to a fresh LB-agar plate from a glycerol stock and were grew overnight at 37˚C. A single colony from the plate was used to inoculate 5 ml fresh LB medium in a 30 ml universal tube, which was then incubated overnight at 37°C with shaking (200 rpm). The overnight culture was diluted 400 fold into 200 ml fresh LB in a 500 ml Erlenmeyer flask, and was cultured at 37°C with shaking (200 rpm) until the OD_{600} reached 0.3 – 0.4. Then the culture was transferred into four pre-cooled 50 ml Falcon tubes, which were kept on ice for 15 min before centrifugation (3,011 g, 7 min, 0°C). Each cell pellet was washed gently with 10 ml pre-cooled sterilised $CaCl_2$ solution (50 mM), and was then incubated on ice for 20 min. Thereafter, the cell pellets were recovered by centrifugation as described above. The cell pellets were washed a second time following the same procedure before being resuspended in 8 ml cold sterilised 50 mM $CaCl_2$+15% (v/v) glycerol (2 ml in each 50 ml Falcon tube). The cells were then combined in one tube and kept on ice for 2 h before being dispensed into pre-cooled 1.5 ml tubes (110 μl in each tube). The cells were stored at – 80˚C.

2.4.2 Heat shock transformation

Chemically competent cells were removed from the − 80˚C freezer and were thawed on ice for 10 min. 0.5 µl – 1 µl plasmid DNA or 10 µl ligation mixture were transferred into 10 µl or 50 µl competent cells respectively, which were gently mixed and incubated on ice for 30 min. The cells were then heat shocked at 42˚C for 60 s (for 10 µl competent cells) or 90 s (for 50 µl competent cells), followed by 3 min incubation on ice. Then the cells were allowed to recover by addition of 0.2 ml LB followed by incubation for 1 h with shaking (200 rpm) at 37°C. For the transformation of a single plasmid or single fragment ligation, 50 µl of recovered cells were plated on a pre-dried agar plate with appropriate antibiotics. For the transformation of multi-fragment ligation or multiple plasmids, all the recovered cells were plated. The plate was then incubated overnight at 37°C.

2.5 *In vivo* gene expression assay

Two different metal concentration terminologies (molar and ppb) were used in different sensor characterisations. The molar concentration was used to facilitate the comparison of sensitivity and output dynamic range with previous studies (Wang et al. 2013; Wang et al. 2014; Wang et al. 2015). Ppb (w/v) concentration was used to aid the comparison with the WHO standard, especially for field testing-related sensor characterisation in Chapter 4.

2.5.1 Assay in liquid culture

The growth conditions for characterising the engineered circuits are described above. Green Fluorescent Protein (GFP) was used as a reporter for all sensor circuits. The *gfp* expression was measured by BMG FLUOstar fluorometry as described above (485 nm for excitation, 520 ± 10 nm for emission, Gain = 1,000). Absorbance (OD_{600}) was also read at the same time to determine the cell density. The fluorometry data were first processed using Omega MARS Data Analysis Software, and then were exported to Microsoft Excel 2013 and GraphPad for data analysis. The medium backgrounds of absorbance and fluorescence were determined from blank wells loaded with LB medium and were subtracted from the readings of other wells. The fluorescence/OD_{600} (Fluo./OD_{600}) at a specific time for a sample culture was determined after subtracting the Fluo./OD_{600} value of the negative control cultures (GFP-free) at the same time. Unless indicated otherwise, the fluorescence/OD_{600} after 5 h growth post initial day dilution and induction was used as the output response of the cells in the steady state when cells were in exponential growth and the steady state assumption for protein expression is applied. Unless indicated otherwise, each sensor was tested with three or more biological replicates. The

sample size (i.e., n) is specified in figure captions. All the data shown are mean values and are based on the plate reader data unless otherwise indicated.

Flow cytometry assays were performed following the plate reader assays. Briefly, the cells from the 96-well plate were transferred and diluted 1 : 100 to another U-bottom 96-well plate (612U96, Fisher Scientific) with PBS (1 ×, with 2 mg/ml Kanamycin to stop translation). The plate was incubated at 4°C for 1 h. Cells were assayed at low flow rate until 10,000 total events were collected using Attune NxT software on an Attune NxT Flow Cytometer (equipped with Attune NxT Autosampler) using 488 nm excitation laser and a 530 nm emission filter with 30 nm bandpass. The flow cytometry data were analysed using Attune NxT software and FlowJo software (v7.6) with an appropriate gate of forward-scattering and side-scattering for all tested cultures.

2.5.2 Agarose gel entrapment-based sensor array assay

The sensor cells were cultured overnight as described above, and then were diluted 200-fold into 10 ml fresh LB medium with appropriate antibiotics in 50 ml Falcon centrifuge tubes. Afterwards, the cells were cultured at 37°C with shaking at 160 rpm until the OD_{600} reached 1.5 (after 4 h 40 min). The cells were concentrated 5-fold in PBS by centrifugation at 3,000 g for 5 min and then resuspension in PBS (K813-500ML, VWR) with 250 μg/ml kanamycin or ampicillin. Strains that failed to reach $OD_{600} \approx 1.5$ were concentrated to the average OD_{600} value of the other strains at this step. Fresh 1.25% (w/v) agarose solution was prepared in PBS (without antibiotics) for each test, and was kept at 55°C in a block heater to prevent solidification before use. The agarose solution was then kept at 42°C for 5 min before mixing with the re-suspended cells at 4 : 1 ratio to re-dilute the cells back to $OD_{600} \approx 1.5$, making the final agarose concentration to 1%, and the added antibiotics diluted to 50 μg/ml. Next the agarose-cell mixture was quickly loaded into a 384-well microplate (781906, Greiner Bio-One) with 15 μl in each well. For the arsenic induction, ddH_2O or $NaAsO_2$ solution was diluted 40-fold into M9 medium (with 0.4% (v/v) glycerol), and 50 μl of the medium-inducer mixture was added to each well. The microplate was then sealed with a sealing membrane (Z380059-1PAK, Sigma-Aldrich) and covered with a plate lid to prevent evaporation. The microplates were incubated at 37°C or 25°C, and fluorescence was measured after 24 h and 40 h.

To test whether the agarose gel-entrapped sensor cell array would give false negative results under lethal antimicrobial conditions, the sensor cell array was tested under the following conditions: 1) 25 μg/ml chloramphenicol, 2) pH = 3.5 (via addition of acetic acid, 338826, Honeywell Research Chemicals), 3) pH = 11.4 (via addition of sodium hydroxide, 71690, Sigma-Aldrich), 4) 1 mM or 3 mM $CuSO_4$.

A variety of media were tested for their abilities to maximise the sensors' responses to arsenate. PBS was replaced by 0.9% NaCl solution in these experiments to resuspend cells and to make agarose solution. M9 medium was optimised by reducing both Na_2HPO_4 and KH_2PO_4 levels to 1/4, 1/16 and 1/64 of the original concentrations, or by substituting the phosphate-based buffer with a Tris-HCl-based buffer (10.992 g/L Tris-HCl, http://2015.igem.org/Team:HKUST-Rice). The pH of M9 (Tris-HCl) medium was adjusted to pH = 7 via addition of 1 M NaOH solution. Since agarose-entrapped cells did not grow when glucose was used as the carbon source (**Appendix Fig. 4.2e,f**), a modified MOPS EZ Rich Defined Medium (M2105, TEKnova) was used where glucose was replaced by glycerol (0.2% (v/v)). The MOPS medium was further optimised by reducing the K_2HPO_4 level to 3/4, 1/2 and 1/4 of its original concentration.

The fluorescent signals of the microplates were measured using a BMG FLUOstar fluorometer, and photographs were taken using a Bio-Rad Gel Doc XR+ system with filter 1 and SYBR Green mode. The images were acquired using Image Lab software at 600 dpi resolution. Each image was first adjusted using Photoshop CS3 software to reduce background by subtracting the brightness of negative controls, and then was adjusted to increase brightness and contrast. To prepare for cell phone imaging, the microplates were placed onto the surface of a Safe Imager™ (S37102, Invitrogen) blue-light transilluminator, and were covered with an amber filter in a dark environment. A cell phone (OPPO X9000) was used to acquire the images with the built-in high dynamic range (HDR) setting, and the images were then adjusted to increase brightness and contrast using Photoshop CS3 software.

2.5.3 Microfluidic encapsulation-based sensor array assay

The experiments in this section were designed and performed by Francesca Volpetti, Ekaterina Petrova and Sebastian J. Maerkl from the Institute of Bioengineering, School of Engineering, Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland.

The microfluidic device is composed of two polydimethylsiloxane (PDMS) layers, a control layer and a flow layer. The molds for the two layers were fabricated using standard photolithography techniques. A 4'' silicon wafer was coated with 30 μm GM1050 SU-8 and 14 μm AZ9260 photoresist for the control and the flow mold respectively. After exposure and development, the flow mold was placed at 160°C for 2 h in order to round the channels. Devices were cast in PDMS of using multilayer soft lithography. PDMS was prepared at a 20 : 1 ratio and spin-coated onto the flow mold at 3,000 rpm for 1 min using a SCS G3P-8 Spin Coater. For the control, PDMS at a ratio of 5 : 1 was poured onto the control mold. Both layers

were cured at 80°C for 30 min and successively aligned. After the alignment, the flow and the control layers were bonded at 80°C for 1.5 h.

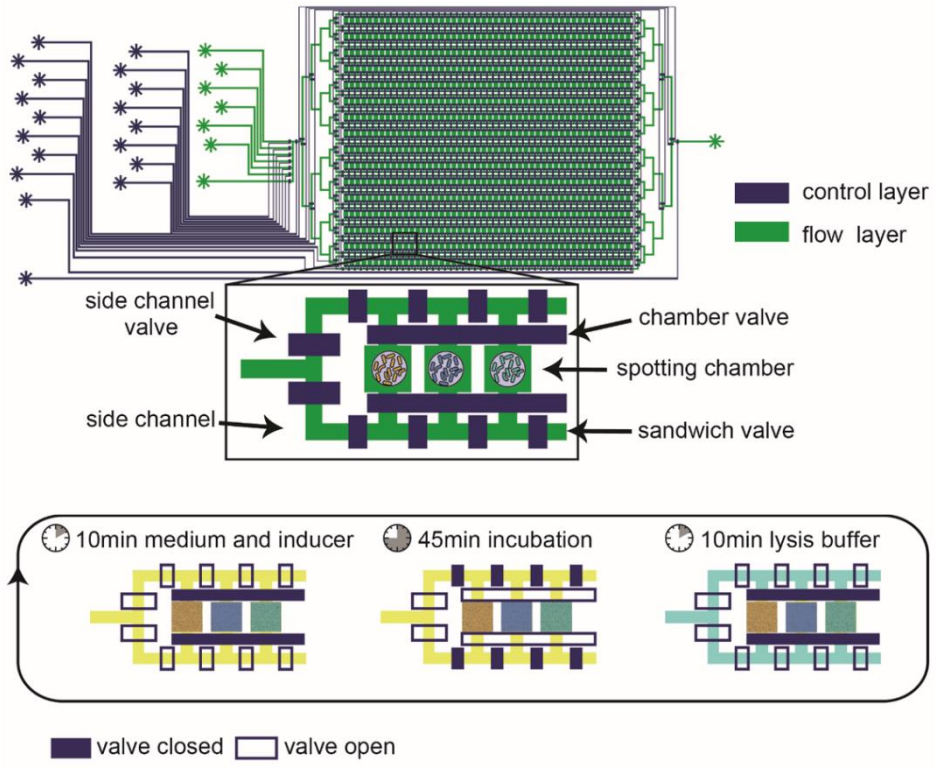


Figure 2.2: Microfluidic biodisplay schematic and use.

Schematic showing the microfluidic biodisplay device and its operation with spotted sensor cells.

The cells were spotted on an epoxysilane (3-Glycidoxypropyl-dimethoxymethylsilane 97%, AC216545000, Acros Organic) functionalised glass slide (**Fig. 2.2**). After overnight growth in LB at 37°C with 200 rpm shaking, the cells were centrifuged for 5 min at 1,781 g and the cell pellet was resuspended in 150 µl LB with 10% glycerol. To generate the pattern, the cells were spotted using a microarray robot (Qarray2, Genetix) and a 1.7 nl delivery volume pin (946MP2B, Arrayit). The microfluidic device was then aligned and bonded to the spotted glass slide for 1 h at 37°C.

After bonding, the control lines of the device were primed with PBS at 5 psi using tygon tubings. Once the control lines were filled the pressure was increased to 10 psi and the flow pressure was set to 2 psi. LabVIEW software was used to facilitate continuous cell culturing and arsenic induction in the device. The culturing program comprises 3 steps: (1) flowing LB medium into the flow channels (keeping the chamber valves closed) for 10 min, (2) incubating the cells (keeping the sandwich valves closed and the chamber valve opened) with the medium

for 45 min, and (3) washing with lysis buffer containing 30 mM of NaOH (06203, Sigma-Aldrich) and 12% SDS (L3771, Sigma-Aldrich) (with chamber valve closed) for 10 min (**Fig. 2.2**). After overnight growth, the inducer $NaAsO_2$ (35000-1L-R, Fluka), diluted in LB, was flowed through the device. A temperature controlled glass plate (H401-NIKON-Ti_SR_Glass / 401_T_CONTROLLER, Okolab) was used to keep the microfluidic device at 37°C during all experiments.

Fluorescent images were acquired using a Nikon ECLIPSE *Ti* automated microscope with a LED fluorescent Excitation System and a Hamamatsu ORCA-Flash 4.0 camera using a 40X magnification objective (SPlan Fluor, ELWD 40X/0.60, ∞/0.2, WD 3.6-2.8, Nikon). A USB fluorescence microscope (AM4113T-GFBW, Dino-Lite) with FITC fluorescent filter was used to acquire fluorescent images at low magnification (10X). Nine subsections of the area of interest were taken and successively stitched together. All the images were adjusted using Photoshop CS3 to increase brightness and contrast. Finally, a mobile phone (iPhone 5) was used to take images of the device after the induction. A band-pass filter (centred at 530 nm with a 40 nm bandwidth) was placed in front of the camera in order to filter out the excitation wavelength and a blue LED was used for illumination. The mobile phone images were adjusted using Photoshop CS3 to reduce background and to increase brightness and contrast.

2.6 Groundwater samples preparation

Groundwater samples were collected from Khulna, Bangladesh. The samples were stored at –20˚C for 6 months. They were filtered through 0.22 μm syringe filters before sending for arsenic quantification by Inductively Coupled Plasma Mass Spectrometry (ICP-MS) at the School of Chemistry in the University of Edinburgh by Dr Lorna Eades, the ICP-MS instrument technician, who also analysed the data. The arsenic quantification of each sample is shown in **Appendix Table 2.5**. For functionality test of the sensor array, the quantified samples were mixed with 2 × M9 or 2 × MOPS (with 1/4 PO_4^{3-} and 0.2% (v/v) glycerol after mixing, see Section 2.5.2) at the ratio of 1 : 1, and were then incubated with the sensor array at 37˚C.

2.7 *In vitro* gene expression assay

Two different metal concentration terminologies (molar and ppb) were used in different sensor characterisations. The molar concentration was used to facilitate the comparison of the sensitivity with the sensors characterised in Chapter 3. Ppb (w/v) concentration was used to aid the comparison with the WHO standard.

2.7.1 Cell-free reactions and data analysis

The *E. coli* S30 Extract System for Circular DNA (L1020, Promega) was used for *in vitro* characterisation. To make the inducer stock solutions, $NaAsO_2$ or $HgCl_2$ was dissolved in nuclease free water and was then filtered using 0.22 μm syringe filters. Polyethylene glycol (PEG) 8000 (6510, OminiPur), PEG 4000 (44273, BDH Chemicals) and Ficoll 400 (F2637, Sigma-Aldrich) were dissolved in nuclease free water to make 50% (w/v) stock solutions and were sterilised by autoclaving.

The cell-free reactions were performed according to the manufacturer's protocol. Each reaction contained 40% (v/v) S30 premix, 10% (v/v) amino acid mix, 30% (v/v) S30 extract, 2% (v/v) inducer. The remaining 18% (v/v) contained DNA template with or without crowding agents and substrates. All reactions, unless otherwise indicated, were prepared in a 384-well plate (781906 or 788096, Greiner Bio-One) on ice, topped with 10 μl of Chill-Out Liquid Wax (CHO1411, Bio-Rad) and the plate was sealed with an air permeable film. Unless otherwise indicated, the plate was incubated and measured continuously by BMG FLUOstar fluorometry at 37˚C (for GFP measurement, 485 nm for excitation, 520 nm for emission, Gain = 1,000; for RFP and mCherry measurement, 584 nm for excitation, 620 ± 10 nm for emission, Gain = 2,000). Details of sample volume, DNA concentration and sample size are specified in figure captions. The fluorometry data were first processed using Omega MARS Data Analysis Software, and were then exported to Microsoft Excel 2013 for data analysis. Unless otherwise indicated, background of fluorescence was subtracted from each well by using the averaged values of negative controls (GFP/RFP/mCherry free). All the data (with replicates) shown are mean values.

2.7.2 Preparation of paper-based cell-free reactions

The papers with wax-printed arrays were gifts from Dr Keith Pardee, and details of the paper preparation can be found in his previous study (Pardee et al. 2014). The wax was melted using a heat block (100˚C) for 3–5 min until the black wax was clearly visible from the opposite side. LacZ substrate 5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-gal) (MB1001, Melford) was dissolved in N,N-dimethylformamide (DMF) (D4551, Sigma-Aldrich) to make 2% (w/v) or 5% (w/v) stock solutions; another substrate, chlorophenol red-β-D-galactopyranoside (CPRG) (59767, Sigma-Aldrich), was dissolved in nuclease free water to make 25% (w/v) stock solutions.

The cell-free reactions were prepared as mentioned above, and 1 or 2 μl reactions were spotted on each circle of the paper at 4˚C. The paper was then incubated in a homemade humidifying

chamber at 37˚C (see details in **Fig. 5.14b**). A cell phone (OPPO X9000) camera with diffused flash was used to take images of the paper during the incubation. To acquire images for the paper with fluorescent output, the paper was placed onto the surface of the Safe Imager™ blue-light transilluminator, and covered with the amber filter in a dark environment. The images were then adjusted to increase brightness and contrast using Microsoft PowerPoint 2013. For the tests shown in **Figure 5.14d–f**, the papers were air dried on a heat block at 55˚C, and were then scanned by a Canon CanoScan Lide 220 Flatbed scanner. The scanned images were aligned using Photoshop CS3 software, and were adjusted to increase brightness and contrast using Microsoft PowerPoint 2013.

2.8 Calculation of sensor detection limit

The limit of detection (LOD) is the lowest analyte concentration likely to be reliably distinguished from the basal signal and at which detection is feasible (Armbruster and Pry 2008). The basal signal can be described as the limit of blank (LOB):

$$\mathrm{LOB} = \mu_B + 1.645 \cdot \mathrm{SD}_B \tag{S1}$$

where LOB is estimated by measuring replicates of a blank sample containing no analyte and calculating the mean (μ_B) and the standard deviation (SD_B). Assuming a Gaussian distribution of the raw analytical signals from blank samples, the LOB represents 95% of the observed values. The LOD is determined by the measured LOB and tested replicates with known low concentration of the analyte. It can be described as below:

$$\mathrm{Output}_{LOD} = \mathrm{LOB} + 1.645 \cdot \mathrm{SD}_L \tag{S2}$$

where SD_L represents the standard deviation of measured replicates with a low concentration of analyte. Assuming a Gaussian distribution again for the low concentration samples, 95% of values will exceed the LOB. When the mean (μ_L) value of the measured replicates is equal to Output_{LoD}, the concentration of the analyte will be the LOD. Therefore, we estimated the LOD of our sensors when:

$$\mu_L - (\mathrm{LOB} + 1.645 \cdot \mathrm{SD}_L) \approx 0 \tag{S3}$$

2.9 Calculation of noise factor and signal to noise ratio

Noise factor (NF) and Signal to Noise Ratio (SNR) were calculated as described in a previous study (Wang et al. 2014). Briefly, the NF is the ratio between the SNR of input and output: $\mathrm{NF} = \mathrm{SNR_{IN}} / \mathrm{SNR_{OUT}}$ (Wang et al. 2014). The SNR is defined by the ratio of the sample's mean fluorescence to its standard deviation at single cell level: $\mathrm{SNR} = \mu / \sigma$ (Wang et al. 2014).

The SNR_{IN} is the SNR at the device input (i.e., sensor output without amplification), while the SNR_{OUT} is the SNR at the device output (i.e., sensor output after amplification). The values were calculated from single cell flow cytometry assays (**Fig. 3.12**, **3.14** and **3.15**).

2.10 Mathematical modelling and data fitting

Biochemical models were developed for individual transcription factor receptor modules to abstract their ligand-dependent dose response behaviour. The ordinary differential equation-based deterministic model was used for accurately modelling the gene regulation and expression across the full input or output range of the sensor systems (Wang et al. 2011a; Wang et al. 2014). For inducible sensor promoters used (*arsR*-P_{arsR} and *merR*-P_{merT} promoters), the promoter P_R (P_{arsR}/P_{merT}) is negatively regulated by its constitutively expressed repressor R (ArsR/MerR) and is responsive to exogenous inducer I (arsenic/mercury) to activate transcription of downstream reporter gene *G*. The output gene expression is modelled by (Alon 2007; Wang et al. 2011a):

$$d[G]/dt = \alpha \cdot k_1 + k_1 \cdot [I]^n/({K_M}^n + [I]^n) - d \cdot [G] \quad \text{(S4)}$$

where $\alpha \cdot k$ is the basal constitutive activity of the promoter, $k \cdot [I]^n/({K_M}^n + [I]^n)$ is the activity due to cooperative transcription activation and $d \cdot [G]$ is the constitutive degradation activity of protein G. K_M and n are the Hill constant and coefficient relating to the promoter-regulator/inducer interaction, k is the maximum expression rate due to induction, α is a constant relating to the promoter basal activity level due to leakage ($0 \leq \alpha < 1$), and d is the degradation rate of G.

The steady state solution of equation S4 is given by:

$$f([I]) = [G]_{SS} = \alpha \cdot k + k \cdot [I]^n/({K_M}^n + [I]^n) \quad \text{(S5)}$$

where $k = k_1/d$ represents the maximum expression level due to induction. Equation S5 gives the reporter protein level at steady state for the inducible promoter P_R and is also the transfer function of P_R. We used this transfer function to fit the characterisation data of the arsenic and mercury inducible promoters and engineered sensor systems using GraphPad software. The best fit parameters and coefficients (with 95% confidence bounds) are listed in **Appendix Table 2.4** and the parameterised transfer functions are plotted in **Figure 3.2b**, **Figure 3.4a**, **Figure 3.6b–c**, **Figure 3.7a**, **Figure 3.8a,c**, **Figure 3.10b,c,e**, **Figure 4.2a**, and **Figure 4.8a–d** respectively against their experimental data. **Appendix Figure 3.2** shows the linear correlation between predicted and experimentally characterised responses of the sensors.

A simple linear mathematical formula was used to model the input-output relationship within the linear amplification dynamic range of the 1-layer amplifier system:

$$y = k \cdot (x - b) \tag{S6}$$

where the slope (also known as amplification gain) $k = \Delta y \,/\, \Delta x$, and b is the constant related to the y-intercept when $x = 0$. The parameterised transfer functions are plotted in **Figure 3.4b** against their experimental data.

Chapter 3. Signal Amplifying Gene Circuits Enable Ultrasensitive Cellular Sensors

Whole-cell biosensors have drawn increasing attention over the last few decades, especially in the rising era of synthetic biology. In contrast to traditional physical and chemical sensors, whole-cell biosensors are renewable, easy-to-manufacture and cost-effective (van der Meer and Belkin 2010; Kim et al. 2018). Accordingly, bacterial cell-based sensors have been studied for various applications, such as environmental monitoring (Stocker et al. 2003; De Mora et al. 2011; Wang et al. 2013; Cerminati et al. 2015; Huang et al. 2015b; Hwang et al. 2016; Kim et al. 2016; Cao et al. 2017; Cayron et al. 2017), clinical diagnosis (Archer et al. 2012; Courbet et al. 2015; Danino et al. 2015; Daeffler et al. 2017; Riglar et al. 2017; Watstein and Styczynski 2018; Mimee et al. 2018) and biotherapy (Duan and March 2010; Saeidi et al. 2011; Gupta et al. 2013; Hwang et al. 2014; Ho et al. 2018), bioproduction (Zhang and Keasling 2011; Zhang et al. 2012), mineral surveying (Cerminati et al. 2011) and landmine clearing (Belkin et al. 2017). Despite their advantages and demonstrated success in the laboratory, very few have been transferred into the market. In addition to biosafety concerns (Dana et al. 2012), cellular sensors are often plagued by their unsatisfactory sensing performance which is insufficient to meet the real-world detection requirements, particularly with regard to low sensitivity (Kim et al. 2016; Shemer et al. 2017; Landry et al. 2018), restricted input/output dynamic ranges (Wang et al. 2014; Wang et al. 2015), and lack of field-deployable and easy-to-interpret user interfaces (Prindle et al. 2012; Volpetti et al. 2017).

Recent advances in synthetic biology have provided numerous solutions for precise gene expression regulation (Ang et al. 2013), such as transcriptional promoter engineering (Ellis et al. 2009; Rajkumar et al. 2013; Wang et al. 2015; Merulla and van der Meer 2016; Chen et al. 2018), translational efficiency tuning (Salis et al. 2009; Wang et al. 2014) and post-translational protein degradation control (Cameron and Collins 2014; Fernandez-Rodriguez and Voigt 2016). While some of these have been applied to improving the performance of genetically encoded biosensors (Hynninen et al. 2010; Nistala et al. 2010; Merulla et al. 2013; Wang et al. 2014; Li et al. 2015; Wang et al. 2015; Brophy and Voigt 2016; Fernandez-Rodriguez and Voigt 2016; Kim et al. 2016; Merulla and van der Meer 2016; Cayron et al. 2017; Chen et al. 2018), most reports only focus on one feature while ignoring others, leading to a trade-off between the different features, in addition to some solutions being case specific. Consequently, the usage of these sensors is inherently restricted, and a complete yet widely

applicable solution for enhancing sensor performance to meet the demands of practical applications is necessary and timely.

To address such challenges, I developed and investigated a novel, modular signal amplifying approach, which can rapidly and drastically improve the detection limit and output amplitude of cell-based biosensors (**Fig. 3.1**). This approach integrates three synergistic signal amplifying strategies by: **(1)** tuning the intracellular receptor protein density to increase sensor sensitivity (Merulla et al. 2013; Wang et al. 2015), **(2)** engineering an ultrasensitive activator-based high-gain transcriptional amplifier to increase output dynamic range (Wang et al. 2014), and **(3)** cascading multiple orthogonal amplifiers in tandem to further boost sensing performance. I first demonstrated these strategies by optimising and engineering ultrasensitive *E. coli*-based sensors for arsenic and mercury water contamination. To address the potential issue of background expression accompanying amplified or sensitive sensors, I proceeded to investigate two approaches based on promoter engineering and post-translational regulation to modulate arsenic sensor background level without compromising its detection limit and maximum output expression (Fernandez-Rodriguez and Voigt 2016; Merulla and van der Meer 2016).

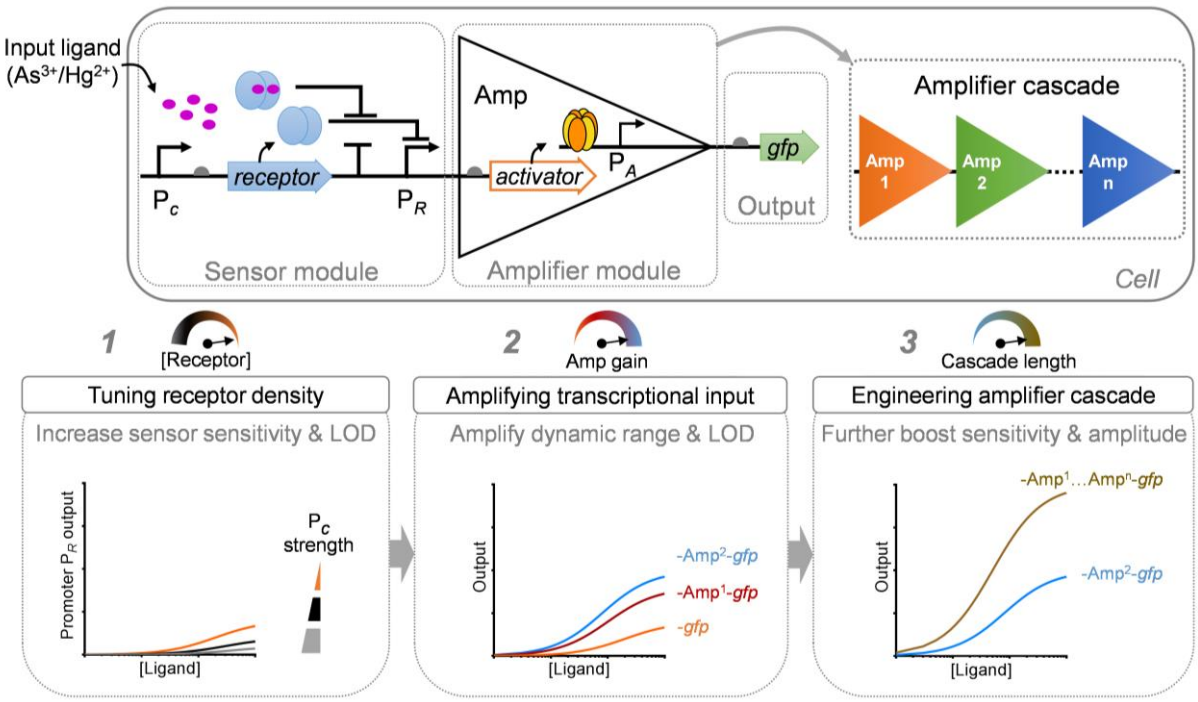


Figure 3.1: Modular multi-layer signal amplification for engineering ultrasensitive transcription-based cellular sensors.

Schematic showing a typical repressor-receptor mediated transcription-based sensor in a bacterial cell. The expression of the small molecule-responsive receptor is driven by a constitutive promoter (P_C). The receptor also acts as a transcriptional repressor, which naturally represses its cognate promoter P_R. When the target molecules (i.e., an input ligand) are present, the receptor binds to its cognate ligand and releases P_R, which then activates the transcription of a downstream gene. The first step (**step 1**) of the signal amplification approach is to tune the intracellular receptor density by varying the strength of P_C. For repressor-based sensor modules, weaker P_C will lead to lower density of the receptor allowing easier release of P_R in the presence of the signal ligand. As illustrated in the lower response diagrams, this results in significant increase of the sensor sensitivity and lowering of the limit of detection (LOD). To expand output dynamic range, a transcriptional amplifier is employed to amplify the transduced transcriptional signal from P_R (**step 2**). The last step (**step 3**) uses a multi-layered amplifier cascade, built from cascaded orthogonal amplifiers, to sequentially amplify the transcriptional input signal flow, and thus further boost the sensor's output readout.

3.1 Increasing sensitivity and output dynamic range by tuning receptor density and employing single-layer transcriptional amplification

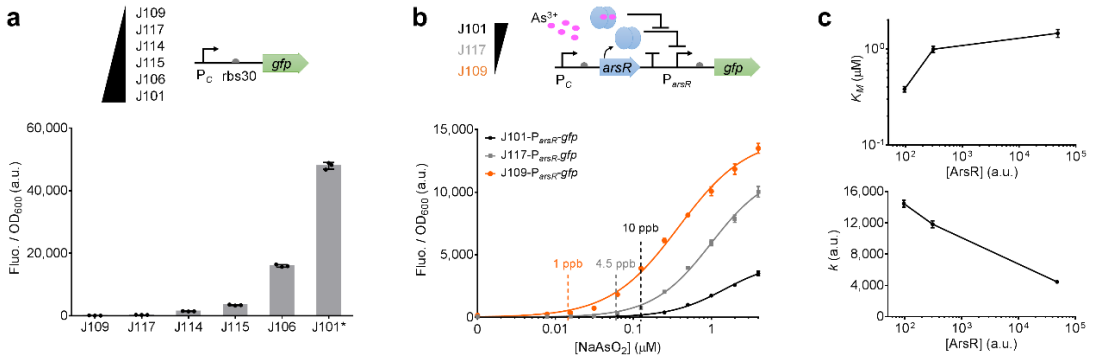


Figure 3.2: Amplifying arsenic sensor by tuning intracellular receptor density.

(**a**) Characterisation of a set of constitutive promoters (P_C) with GFP as output. J101*: rbs32 was used for the J101 promoter. Error bars, s.d. (n = 3). (**b**) Tuning arsenic receptor ArsR's intracellular density by varying the strength of constitutive promoter P_C, showing that a weaker promoter can improve both sensitivity and output readout of arsenic sensors. Error bars, s.d. (n = 4). (**c**) The Hill constant (K_M) and the maximum output (k) of the sensor's dose–response against the relevant intracellular ArsR concentrations (i.e., [ArsR], defined by the strength of P_C, which is measured by the output fluorescence level shown in **a**). Error bars, 95% confidence intervals. a.u., arbitrary units (Fluo./OD_{600}).

I first tested the signal amplifying approach on a previously studied arsenic sensor (i.e., J101-*arsR*-P_{arsR}-*gfp*) (Wang et al. 2013). This sensor (**Fig. 3.2b**) has a constitutive promoter (J101) that drives the expression of an arsenic receptor ArsR, which would de-repress its cognate promoter P_{arsR} upon arsenite binding and trigger the expression of a reporter gene, *gfp*. Unfortunately, this sensor can only sense > 10 ppb arsenic (i.e., 0.133 μM), so it would not meet the requirement for real applications (10 ppb is the safety upper limit set by Environmental Protection Agency (EPA) and the World Health Organisation (WHO) for arsenic in drinking water).

To perform the first optimisation strategy by tuning the intracellular receptor density, I characterised six constitutive promoters of varying strengths (**Fig. 3.2a**), and chose two weak promoters (i.e., J117 and J109) to replace J101 in the arsenic sensor. The sensors were then compared under various arsenic ($NaAsO_2$) induction conditions (**Fig. 3.2b**). The results showed that the weaker the promoter (i.e., the lower the ArsR receptor concentration), the more sensitive and higher the dynamic range of the sensor. That said, there is a limitation on the promoter strength as extremely weak promoter may lead to high basal expression and low

output dynamic range. Notably, the arsenic sensor with the weakest promoter (J109) could sense 1 ppb arsenic (with output 355.3 ± 37.2 a.u. compared to 170.9 ± 25.5 at 0 ppb arsenic induction), which is 10 times lower than the original sensor.

I fitted the sensors' response curves to a Hill function-based biochemical model to describe their input-output relationships (see Materials and Methods, **Appendix Table 2.4**). In this case, the Hill constant K_M is the inducer concentration that provokes half-maximal activation of a sensor, and k is a sensor's maximum output expression level; therefore, K_M is negatively correlated with sensitivity, while k is positively correlated with output. Here K_M showed a 4-fold decrease and k showed a 3-fold increase from high to low ArsR levels (**Fig. 3.2c**), confirming that the arsenic sensor's sensitivity and output were increased while the ArsR intracellular density was decreased.

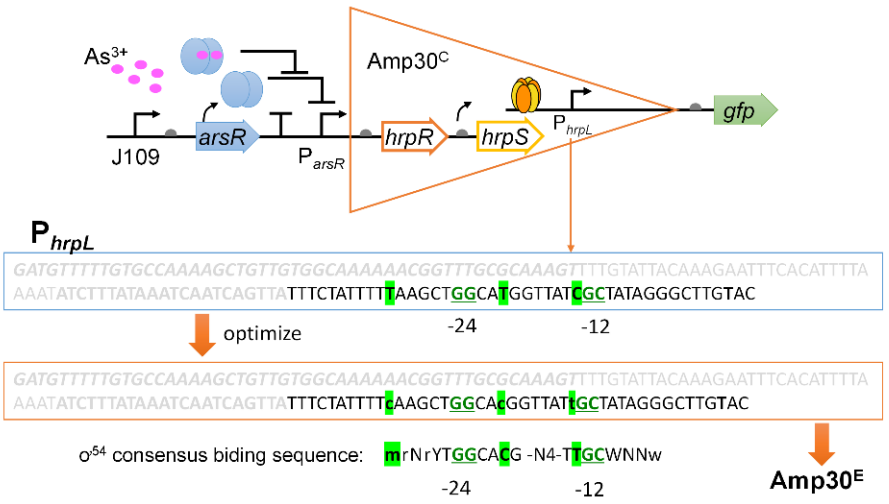


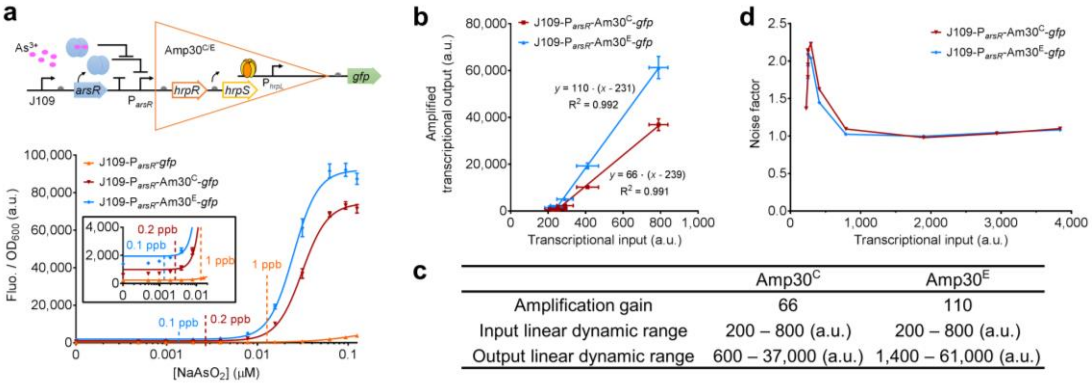
Figure 3.3: Design of the amplifier Amp30^C and its optimised version Amp30^E.

Schematic diagram showing the optimisation of amplifier Amp30^C. Amp30^C is built by expressing in an operon the cooperative activator proteins, HrpR and HrpS, whose high order functional forms synergistically activate the downstream tightly-controlled σ^{54}-dependent P_{hrpL} promoter. We hypothesised that by changing the σ^{54} binding sequence to a consensus sequence (mrNrYT**GG**CACG-N4-TT**GC**WNNw) (Barrios et al. 1999) in P_{hrpL}, we would increase the promoter's transcriptional efficiency, and thereby improve the amplifier's maximum output level. The σ^{54} binding –24 and –12 sites are in bold green, and mutated nucleotides are highlighted in green.

To further improve the output expression, I proceeded to introduce a transcriptional amplifier between the sensor module and the reporter. A recently developed amplifier, Amp30^C, can amplify transcriptional inputs and increase the output, but its amplification at low input is limited (Wang et al. 2014). To achieve better amplification, I retrofitted Amp30^C to generate

amplifier Amp30^{E} by adopting the consensus σ^{54} promoter sequence for its output P_{hrpL} promoter (**Fig. 3.3**), and compared them using the most sensitive arsenic sensor module (i.e., J109-*arsR*-P_{arsR}). **Figure 3.4a** shows that both amplifiers increased the output expression, and the retrofitted Amp30^{E} performed better than Amp30^{C}. Notably, the sensor comprising Amp30^{E} could sense 0.1 ppb of arsenic (100-fold improvement of detection limit), and increased the output expression by 440-fold compared to the original arsenic sensor. I used a simple linear mathematical formula to model the input-output relationship within a linear amplification range for both amplifiers. The results showed a higher amplification gain for Amp30^{E} (gain = 110) than Amp30^{C} (gain = 66) (see Materials and Methods, **Fig. 3.4b,c**). This further confirmed that the amplifier optimisation improved amplification capacity.

To analyse the signal fidelity of the amplifiers, I calculated their noise factors (NF) across varying transcriptional signal inputs (see Materials and Methods). Low NFs (NF < 2.5) were observed for both Amp30^{C} and Amp30^{E} (**Fig. 3.4d**), indicating that no significant noise was introduced by the signal amplification and that the optimised amplifier did not further increase the noise.



	Amp30^{C}	Amp30^{E}
Amplification gain	66	110
Input linear dynamic range	200 – 800 (a.u.)	200 – 800 (a.u.)
Output linear dynamic range	600 – 37,000 (a.u.)	1,400 – 61,000 (a.u.)

Figure 3.4: Amplifying arsenic sensor by employing single-layer transcriptional amplification.

(**a**) Further amplifying the sensitivity and output of arsenic sensors by using a transcriptional amplifier (Amp30^{C} and Amp30^{E}), showing improved amplification of optimised amplifier Amp30^{E}. The orange curve shown in **a** is the same construct as the orange curve in **Fig. 3.2b**. Error bars, s.d. (n = 4). (**b**) A scatter plot showing the linear relationships between the non-saturated transcriptional inputs (i.e., output levels of the J109-P_{arsR}-*gfp* sensor characterised in **a**) and the amplified outputs of Amp30^{C} and Amp30^{E}. Error bars, s.d. (n = 4). (**c**) Summary of performance characteristics of Amp30^{C} and Amp30^{E}. (**d**) Device noise characteristics of Amp30^{C} and Amp30^{E}. Transcriptional input is defined by the output fluorescence level from the sensor without amplifiers, which is J109-P_{arsR}-*gfp* shown in **a**. a.u., arbitrary units (Fluo./OD_{600}).

3.2 Sequential cascaded amplification further boosts sensor sensitivity and output amplitude

In principle, the transduced transcriptional signal of the sensor could be sequentially amplified by multi-layer coupled amplifiers. To test this hypothesis, I engineered alternative high-gain transcriptional amplifiers using a phage activator RinA_p80α (from *Staphylococcal aureus* phage 80α) (Quiles-Puchalt et al. 2013) and an extracytoplasmic function (ECF) sigma factor ECF11_987 (from *Vibrio parahaemolyticus*) (Rhodius et al. 2013) (**Fig. 3.5**). An orthogonality characterisation was performed, showing no cross activation among the RinA, ECF11 and the HrpRS-based amplifiers (**Fig. 3.6a**) and hence their potential to be utilised in one system.

It should be noted that, to realise effective signal amplification, a functional amplifier cascade requires the input-output profile of different coupled amplifiers to be matched (Nielsen et al. 2016), and the output signal of an amplifier should be higher than its input signal. Accordingly, I built a library of the RinA/ECF-based amplifiers of different input/output profiles and amplification capacities (**Fig. 3.5**).

From this library, I selected Amp31E11^{A} as the second layer to be connected to the first layer Amp30^{E} to build a 2-layer amplifier cascade. To minimise sensor's basal leakage level and cellular burden, I connected the amplifier cascade to a tight input sensor module (i.e., J117-*arsR*-P$_{arsR}$) in a low copy number plasmid (i.e., pSB4A3, ~ 5 copies; previous sensor was in pSB3K3, 10 ~ 12 copies) (Shetty et al. 2008; Liu et al. 2018a). It was found that the 2-layered amplifier (orange curve in **Fig. 3.6b**) achieved higher signal output than Amp30^{E} alone (black curve in **Fig. 3.6b**) at low arsenic induction levels. Additionally, in contrast to the single-layer amplification using Amp30^{E} alone, the 2-layered amplifier also improved the detection limit by 4-fold. Further investigation (by characterising the sensor with Amp30^{E} alone in pSB4A3) showed that such sensing performance improvement is largely owing to the 2-step amplification whereas the low-copy plasmid may only contribute slightly to the detection limit improvement (grey dashed curve in **Fig. 3.6b**).

I then appended a third amplifier, Amp33RinA, to the 2-layered amplifier. Surprisingly this 3-layer amplified sensor exhibited a poorer dose-response (red curve in **Fig. 3.6c**) than the 2-layer amplified one, possibly due to the amplifiers' incompatible input/output profiles caused by the load of the amplifier cascade on the host or competition for usage of limited cellular resources. To rescue the amplifying function of the third layer, I removed the final amplifier's output promoter along with the reporter to a high copy number plasmid (i.e., pSB1K3, > 100 copies) (**Fig. 3.6c**). I hypothesised that the high copy number plasmid could improve the

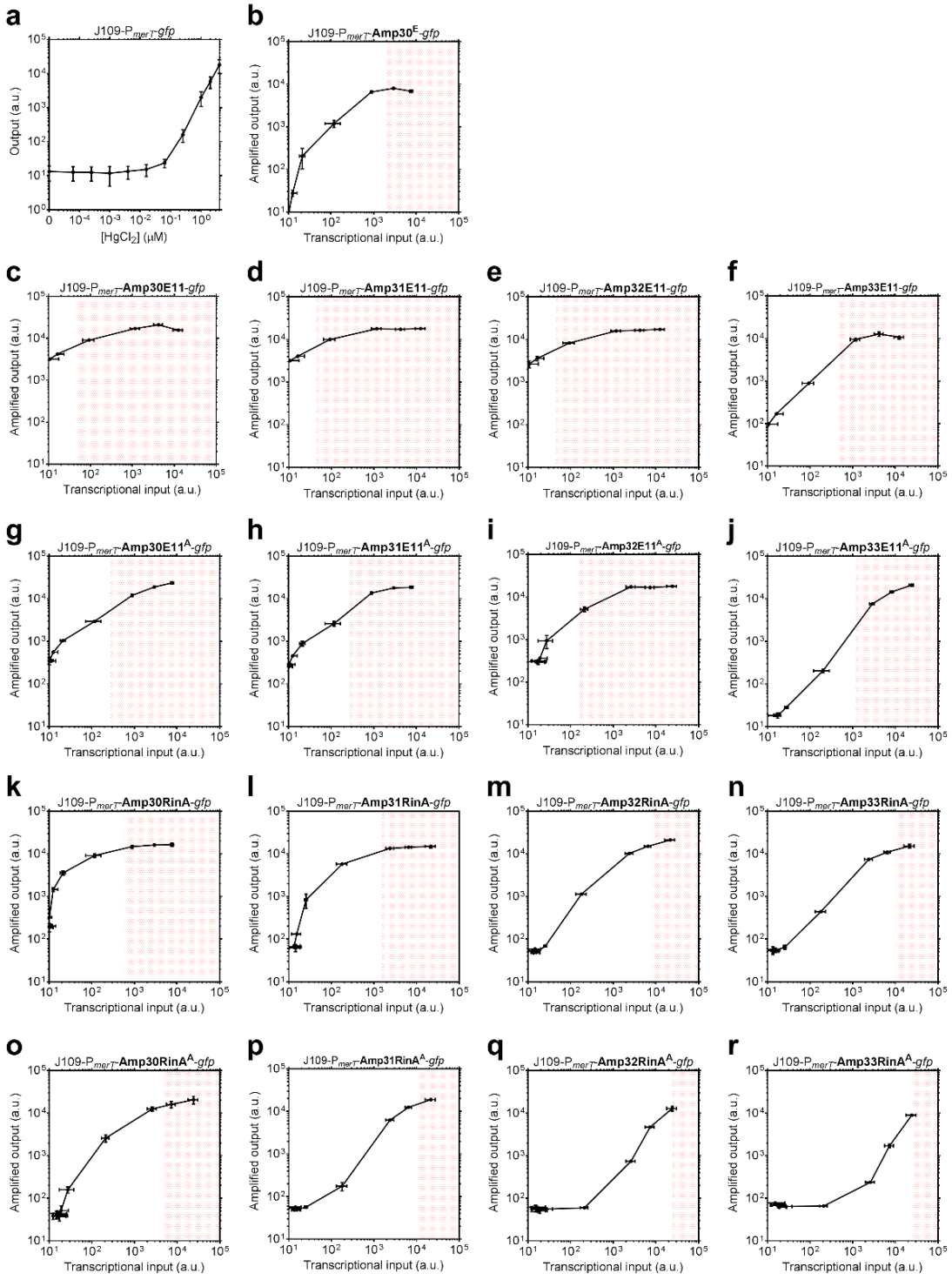


Figure 3.5: Characterised input-output dose responses for a library of engineered transcriptional amplifiers.

The library comprises RinA/ECF-based amplifiers with different expression levels of the activators, which are achieved by varying ribosome binding site and adding degradation tags to the activators. A mercury responsive promoter P_{merT} was characterised under various mercury induction (**a**). Using this promoter as the transcriptional input, HrpRS-based amplifier (**b**, Amp30^E), ECF11_987-based

activator-promoter binding efficiency (Potvin-Trottier et al. 2016), amplify the output amplitude, and thereby lead to enhanced amplification. The new 3-layered amplifier with the output on the high copy number plasmid (blue curve in **Fig. 3.6c**) dramatically amplified the output expression, and displayed a clearly better performance than the 2-layered amplifier. Moreover, the high copy number plasmid further improved the detection limit (from 1.5 ppb to 0.8 ppb), suggesting that the high copy number plasmid may also contribute to shifting or expanding the amplifier's input signal range.

I fitted the sensor dose-response curves to the same aforementioned Hill function model, and found that the Hill constant K_M was decreased when the number of amplifiers was increased in the cascade (**Fig. 3.6d**, see Materials and Methods, **Appendix Table 2.4**). Meanwhile *k* was increased, meaning the sensor output expression was stepped up when the transcriptional signal was sequentially amplified by multi-layered amplifiers. These findings confirmed that multi-layered amplifiers can further boost sensor sensitivity and output amplitude.

amplifiers (**c**–**j**) and RinA_p80α-based amplifiers (**k**–**r**) were engineered and tested individually. Letters in bold indicate amplifier's name. Details of each amplifier are listed in **Appendix Table 2.2**. Red shaded regions indicate the input ranges that caused toxicity to host cells (as judged by a fall in the 5 h OD_{600} to 75% or lower of uninduced). Transcriptional input is defined by the output fluorescence level from the sensor without amplifiers, which is J109-P_{merT}-*gfp* shown in **a**. Output and amplified output are the output fluorescence level of the sensor without or with amplifiers. All data shown are median values from flow cytometry assay. Error bars, s.d. (**a**, n = 15; **b**–**r**, n = 3). a.u., arbitrary units.

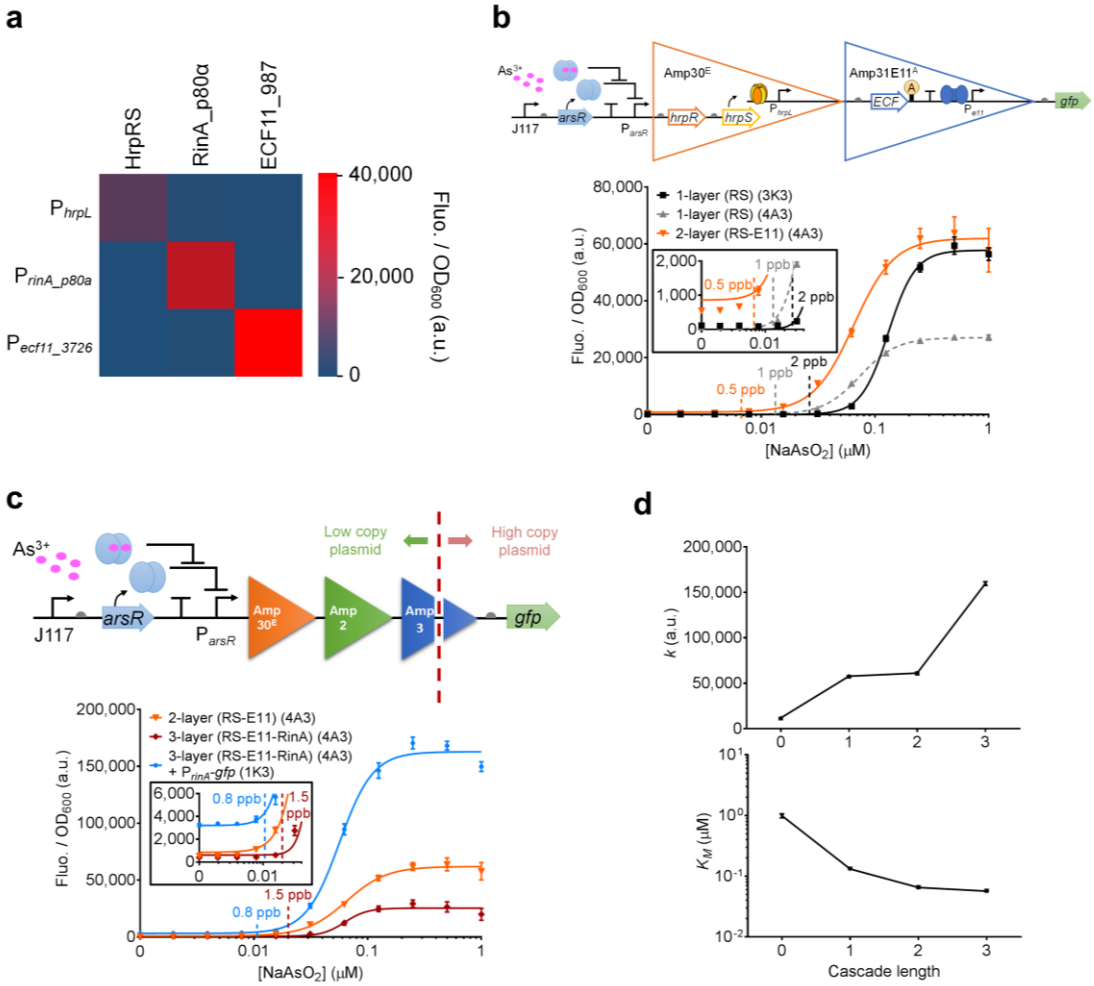


Figure 3.6: Sequential cascaded amplification further boosts sensor sensitivity and output amplitude.

(**a**) Orthogonality characterisation of the three transcriptional amplifiers showing no mutual cross talk. 0.25 μM $HgCl_2$ was used for the assay. Data shown are mean values of three biological repeats, and were collected 8 h post induction and incubation. 5 h data were not shown due to negative values. (**b**) Design and characterisation of an arsenic sensor with 2-layer amplifier cascade, showing improved detection limit and output readout. 1-layer arsenic sensors are J117-P_{arsR}-*gfp* carried by pSB3K3 (3K3) or pSB4A3 (4A3). Ⓐ represents ASV degradation tag. Error bars, s.d. (n = 4). (**c**) Design and characterisation of an arsenic sensor with 3-layered amplification implemented on two plasmids, showing the output can be further boosted with additional layer of amplification. Error bars, s.d. (n = 4). (**d**) Plots showing the Hill constant (K_M) and the maximum output (k) of the sensor's dose-response. Error bars, 95% confidence intervals. The detailed configurations and plasmid maps of the single and multi-layered amplifiers are shown in **Appendix Figure 3.1a,c.** RS, HrpRS-based amplifier. E11, ECF11_987-based amplifier. RinA, RinA_p80α-based amplifier. a.u., arbitrary units (Fluo./OD_{600}).

3.3 Signal amplification is modular without compromising sensor specificity

To verify whether the multi-step amplification is modular, I swapped the order of RinA and ECF11-based amplifiers in the previous amplifier cascades (**Fig. 3.7**). I found that the alternative 2-layered and 3-layered amplifiers similarly improved the detection limit and the output expression (**Fig. 3.7a**), as also confirmed by the model fitted parameters of their dose-response curves (**Fig. 3.7b**). These results suggest that the multi-step amplification was achieved independently from the type of amplifiers used.

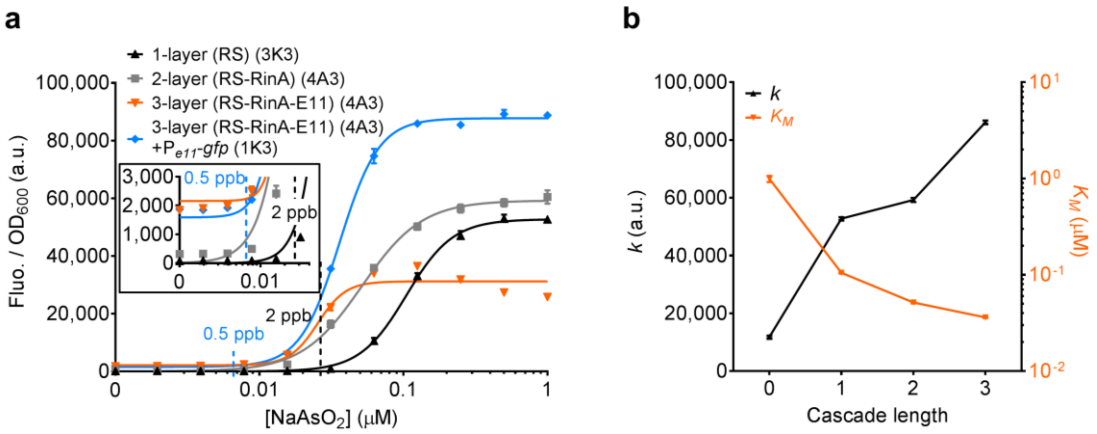


Figure 3.7: Alternative multi-layered amplifiers further boost sensor sensitivity and output amplitude.

(**a**) Characterisation of arsenic sensors with alternative multi-layered amplifiers, demonstrating the modularity of the sequential cascaded amplification. Error bars, s.d. (n = 3). (**b**) Plot showing the Hill constant (K_M) and the maximum output (k) of the sensor against the length of the amplifier cascade (numbers indicating the number of layered amplifiers used). Error bars, 95% confidence intervals. a.u., arbitrary units.

To test the modularity of the entire set of signal amplifying strategies, I next applied it to a different sensor system, a mercury responsive MerR receptor-based sensor. Unlike ArsR, MerR is a repressor-activator for its cognate promoter P$_{merT}$ (Chang et al. 2015). I selected a previously studied mercury sensor (J115-*merR*-P$_{merT}$-*gfp*) (Wang et al. 2013) to perform the optimisation. In this sensor, a constitutive promoter (J115) drives the expression of MerR which regulates the activation of P$_{merT}$ for output reporter *gfp* expression. I first replaced J115 with weaker ones, J114 and J109. As expected, this receptor concentration tuning improved the sensor's detection limit (from 50 ppb to 0.3 ppb) and the output dynamic range (**Fig. 3.8a**). I next introduced the 1-layered, 2-layered, and 3-layered amplifiers to the most sensitive mercury sensor obtained from receptor density tuning (i.e., J109-*merR*-P$_{merT}$) (**Fig. 3.8b,c**).

The results show that all amplification variants significantly boosted the sensor's detection limit and output. Compared to the original mercury sensor, the signal amplifying approach improved the detection limit 5,000-fold, and increased the output amplitude 200-fold at 2 ppb (EPA safety limit) mercury induction. Similarly, the sensor's fitted dose-response showed that its Hill constant K_M was decreased dramatically (> 3 orders of magnitude) as the number of amplifiers was increased in the amplifier cascade (**Fig. 3.8d**). This further demonstrated that ultrasensitive sensors could be achieved by using synergistic multi-layered amplification, and that the signal amplifying approach can be applied to different types of sensor.

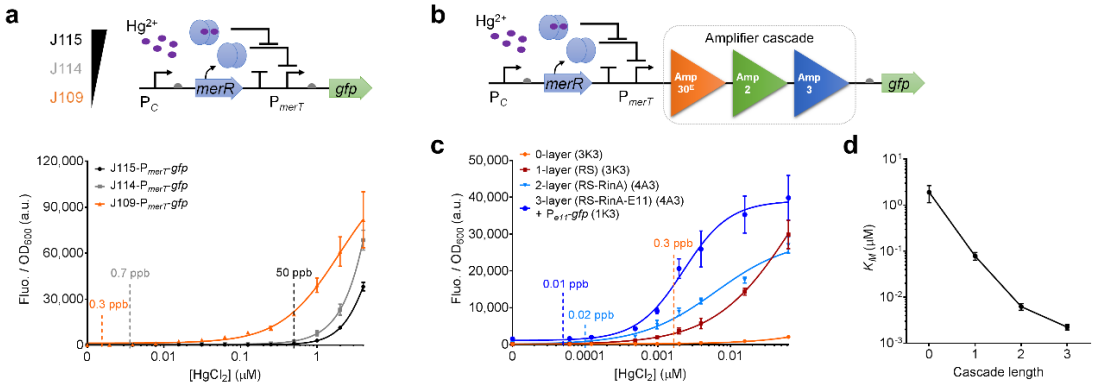


Figure 3.8: Synergistic multi-layered amplification enables ultrasensitive sensors for mercury.

(**a**) Design of the mercury-responsive sensor and tuning of MerR receptor intracellular concentration by varying the strength of its constitutive driving promoter P_C. Data shown were collected 6 h post induction and incubation. Error bars, s.d. (n = 3). (**b**) Design of mercury sensor with multi-layered amplification. Detailed configurations of the multi-layered amplifiers are shown in **Appendix Figure 3.1b**. (**c**) Mercury sensor's dose-response by using different number of layers of amplification. Error bars, s.d. (n = 3). (**d**) Hill constant (K_M) of the sensors' fitted dose-response against different amplifier cascade lengths. Error bars, 95% confidence intervals. a.u., arbitrary units.

Additionally, I tested the sensors' specificity against a wide range of metals that could be potential water contaminants (**Fig. 3.9**). The results show that the amplified sensors exhibited the same high specificity as their original non-amplified arsenic and mercury sensors (Wang et al. 2013). Therefore, the signal amplification enhanced sensing performance without compromising the sensor specificity.

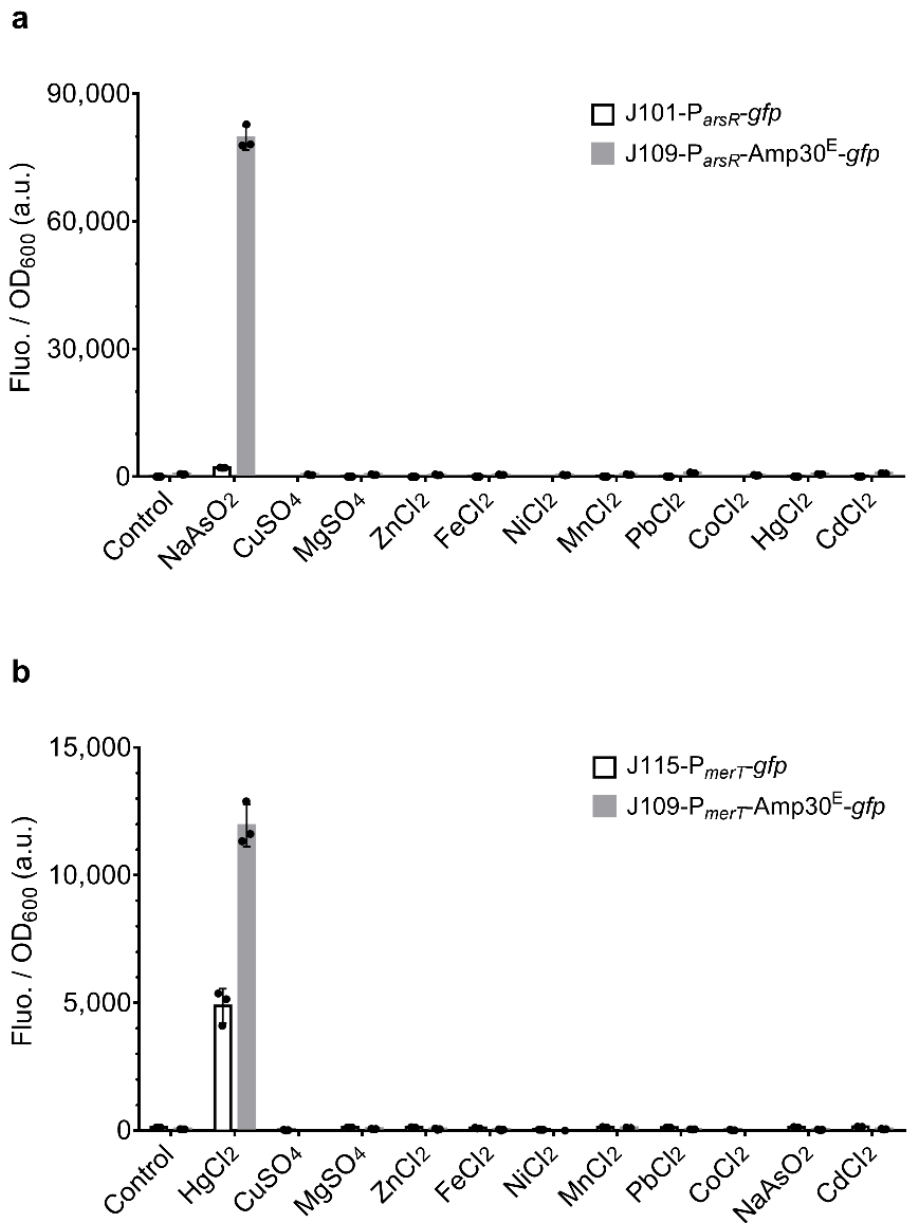


Figure 3.9: Sensor specificity characterisation without and with transcriptional amplification.

(**a**) Specificity test of arsenic sensors. 2 μM or 0.25 μM of $NaAsO_2$ was used to induce the original and amplified arsenic sensor respectively. $HgCl_2$ was added at 2 μM, $CdCl_2$ at 1 μM, and all other metals at 100 μM. (**b**) Specificity test of mercury sensors. 1 μM or 0.125 μM of $HgCl_2$ was used to induce the original and amplified mercury sensor respectively. $NaAsO_2$ was added at 2 μM, $CdCl_2$ at 1 μM, and all other metals at 100 μM. Error bars, s.d. (n = 3). a.u., arbitrary units.

3.4 Tuning sensor's background leakage and output dynamic range

I noticed detectable basal expression (leakage) from the characterisation of 1- and 2-layered amplifiers coupled to the most sensitive arsenic sensor module (i.e., J109-*arsR*-P_{arsR}) (**Fig. 3.10b,c**). Though such leakage may be ignored under certain application circumstances, it could cause issues in other settings, for example, **(1)** sensitive enzyme-based colorimetric output may easily saturate with high basal expression, restricting titrimetric analysis (Wackwitz et al. 2008); **(2)** narrowing down output dynamic range of the downstream reporter expression (Nielsen et al. 2016); **(3)** causing non-stringent side effects for sensor outputs that have therapeutic killing functions.

To reduce sensor leakage, I tested and integrated two different approaches. For the first approach, I inserted an extra ArsR binding site (ABS) downstream of P_{arsR} to create a 'roadblocking' effect (Merulla and van der Meer 2016). Additionally, by tuning the distance between the two ABS (**Fig. 3.10a**, **Appendix Table 2.1**), the sensor's leakiness and sensitivity can be adjusted while maintaining the maximum output, leading to the modification of the sensor's input and output dynamic ranges (**Fig. 3.10b,c**).

The second approach uses protease-based post-transcriptional degradation regulation (Fernandez-Rodriguez and Voigt 2016). I first showed that adding a protein degradation tag (AAV) to the reporter protein reduced the output basal expression, but this also significantly lowered the sensor's sensitivity and output level (grey curve in **Fig. 3.10e**). To obtain low basal level without sacrificing the high output, I next incorporated the sensor into a TEV protease-based reporter protein degradation control system, and used an ECF11-based amplifier to enhance the protease expression; the TEV protease can cleave the linker between the expressed reporter GFP and its fused AAV tag (**Fig. 3.10d**). To optimise the system, I kept the original P_{arsR} in the sensor module to maintain the sensitivity, but used the engineered P_{arsR} containing double ABSs to achieve tight expression control of the TEV protease. The characterisation shows that this design fully protected GFP reporter from degradation at high arsenic induction levels, while achieving significantly lower basal expression through continuous degradation of the reporter GFP at low arsenic induction levels (orange curve in **Fig. 3.10e**). Notably, this integrated approach lowered the detection limit from 5 ppb to 0.2 ppb, which is similar to that of the original leaky amplified arsenic sensor. In summary, this hybrid regulation system was sufficient to reduce the sensor's basal output expression while also maintaining both sensor output amplitude and sensitivity, leading to expanded output dynamic range.

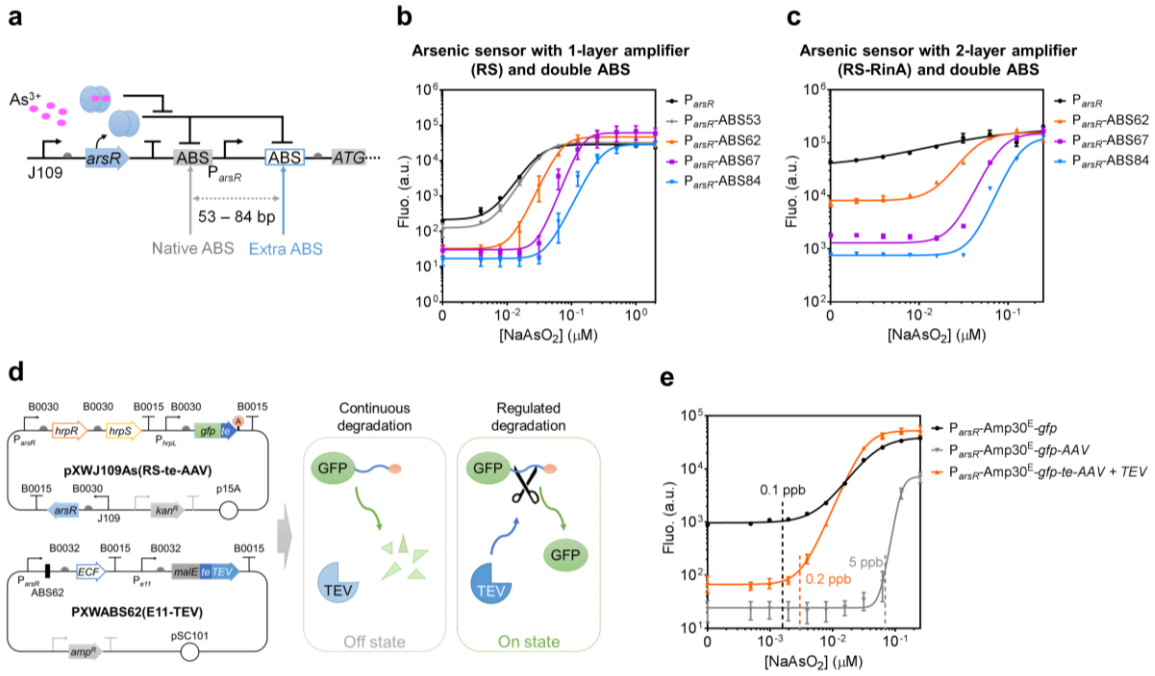


Figure 3.10: Tuning the sensor background and output dynamic range via promoter engineering and reporter degradation regulation.

(**a**) Engineering arsenic promoters by inserting an extra ArsR binding site (ABS) downstream of P_{arsR}. The distance between the two ABSs varies between 53 bp and 84 bp. (**b** and **c**) Dose-responses of engineered arsenic sensors as in **a**, showing that double ABSs can reduce the background output of arsenic sensor (J109-P_{arsR} sensor module) with 1 or 2-layered amplifier, while maintaining the maximum output levels. P_{arsR}-ABS53/62/67/84, engineered P_{arsR} with a second ABS 53/62/67/84 bp downstream of the native ABS. Detailed sequences are listed in **Appendix Table 2.1**. (**d**) Schematic showing protease-mediated regulation of the background and output dynamic range for an arsenic sensor. Ⓐ and ⬭ represent AAV degradation tag. Off state: when there is no $NaAsO_2$ induction. On state: when there is $NaAsO_2$ induction. (**e**) Dose-responses of the arsenic sensor as in **d** 8 h post induction and incubation. Black curve represents the sensor without AAV tag. Grey curve represents the sensor with AAV tag not cleavable by TEV protease. Orange curve represents the genetic circuit shown in **d**, in which the AAV tag can be cleaved off by TEV protease. All data shown are median values from flow cytometry assay. Error bars, s.d. (n = 3). a.u., arbitrary units.

3.5 Discussion

Table 3.1: Summary of signal amplified sensors' LOD and output readout

Sensor	Target	LOD (ppb)	Output (a.u.)*	Tested in
J101-P_{arsR}-*gfp*	$NaAsO_2$	10.00	207.52	Fig. 3.2b
J117-P_{arsR}-*gfp*	$NaAsO_2$	4.50	1,051.98	
J109-P_{arsR}-*gfp*	$NaAsO_2$	1.00	3,783.13	
J109-P_{arsR}-*gfp*	$NaAsO_2$	1.00	4,025.88	Fig. 3.4a
J109-P_{arsR}-Am30^{C}-*gfp*	$NaAsO_2$	0.20	73,629.52	
J109-P_{arsR}-Am30^{E}-*gfp*	$NaAsO_2$	0.10	91,709.65	
1-layer (RS) (3K3)	$NaAsO_2$	2.00	29,550.35	Fig. 3.6b and c
1-layer (RS) (4A3)	$NaAsO_2$	1.00	22,441.42	
2-layer (RS-E11) (4A3)	$NaAsO_2$	0.50	52,928.34	
3-layer (RS-E11-RinA) (4A3)	$NaAsO_2$	1.50	24,475.79	
3-layer (RS-E11-RinA) (4A3) + P_{rinA}-*gfp* (1K3)	$NaAsO_2$	0.80	150,999.20	
1-layer (RS) (3K3)	$NaAsO_2$	2.00	34,561.86	Fig. 3.7
2-layer (RS-RinA) (4A3)	$NaAsO_2$	0.50	52,012.98	
3-layer (RS-RinA-E11) (4A3)	$NaAsO_2$	0.50	31,202.55	
3-layer (RS-RinA-E11) (4A3) + P_{e11}-*gfp* (1K3)	$NaAsO_2$	0.50	86,413.18	
J115-P_{merT}-*gfp*	$HgCl_2$	50.00	238.35	Fig. 3.8a
J114-P_{merT}-*gfp*	$HgCl_2$	0.70	538.49	
J109-P_{merT}-*gfp*	$HgCl_2$	0.30	2,127.57	
0-layer (3K3)	$HgCl_2$	0.30	703.02	Fig. 3.8b
1-layer (RS) (3K3)	$HgCl_2$	0.02	10,884.99	
2-layer (RS-RinA) (4A3)	$HgCl_2$	0.02	16,563.86	
3-layer (RS-RinA-E11) (4A3) + P_{e11}-*gfp* (1K3)	$HgCl_2$	0.01	34,701.04	

*: data shown are values from fitting curves at 10 ppb $NaAsO_2$ or 2 ppb $HgCl_2$ induction.

I developed a modular cascaded signal amplifying approach and combined it with basal background tuning approaches to provide an integrated solution for improving the sensitivity and output dynamic range of cell-based sensors. Using the signal amplifying approach, I drastically increased the sensitivity and output readout of the two exemplar arsenic and mercury sensors, and achieved detection limits of ≤ 0.1 ppb for arsenic and ≤ 0.01 ppb for mercury respectively. Response performance of amplified sensors are summarised in **Table 3.1**. Owing to its modularity and simplicity, the method presented here can be readily applied to improve the sensing performance of many other cell-based and potentially cell-free

genetically-encoded sensors (Pardee et al. 2014; Rampley et al. 2017) to meet their detection requirements for a broad range of real world applications, including environmental assessment (Stocker et al. 2003; De Mora et al. 2011; Wang et al. 2013; Cao et al. 2017), disease diagnosis (Courbet et al. 2015; Danino et al. 2015; Pardee et al. 2016a; Riglar et al. 2017; Wen et al. 2017; Cevenini et al. 2018; Watstein and Styczynski 2018; Mimee et al. 2018), bioproduction (Zhang and Keasling 2011; Zhang et al. 2012), mining and detection of landmines (Cerminati et al. 2011; Belkin et al. 2017).

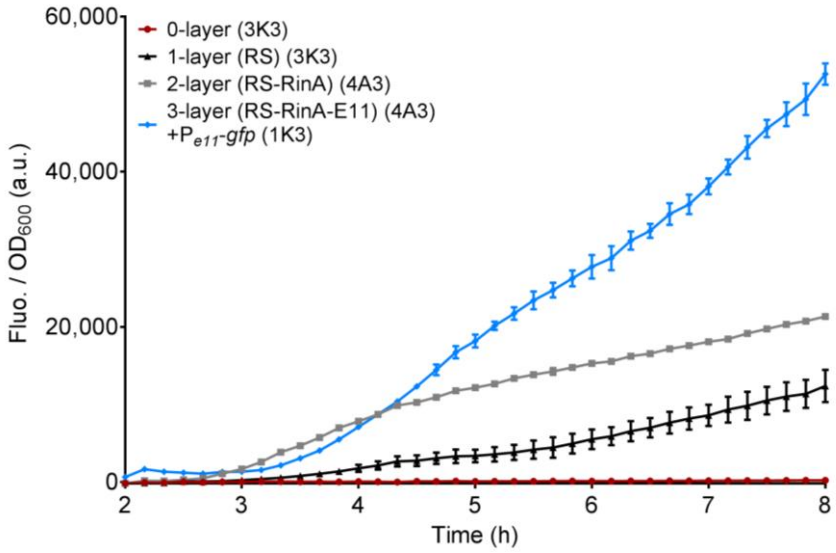


Figure 3.11: Dynamic output responses of arsenic sensors.

Dynamic output responses of arsenic sensors responding to 10 ppb As of $NaAsO_2$ induction, showing that no significant time delay has been caused by the introduction of amplifiers. Error bars, s.d. (n = 3). a.u., arbitrary units.

I showed for the first time that multiple layers of transcriptional amplifiers can in-tandem amplify a transduced sensor signal *in vivo*. Notably, these engineered high-gain amplifiers and amplifier cascades did not introduce obvious response delay (**Fig. 3.11**), and did not constitute significant sources of noise during the signal amplification stage (**Fig. 3.12**), which are important for modulating biological signals due to their inherent slow dynamics and noisy characteristics. Moreover, the largest constructs (containing 23 bioparts carried on two plasmids) did not show any notable toxicity to host cells when induced by target pollutants within their WHO/EPA safety limit levels (**Fig. 3.13**). This indicates that the metabolic load of the amplifier constructs was compliant to common *E. coli* host strains which should facilitate their adoption in many application scenarios.

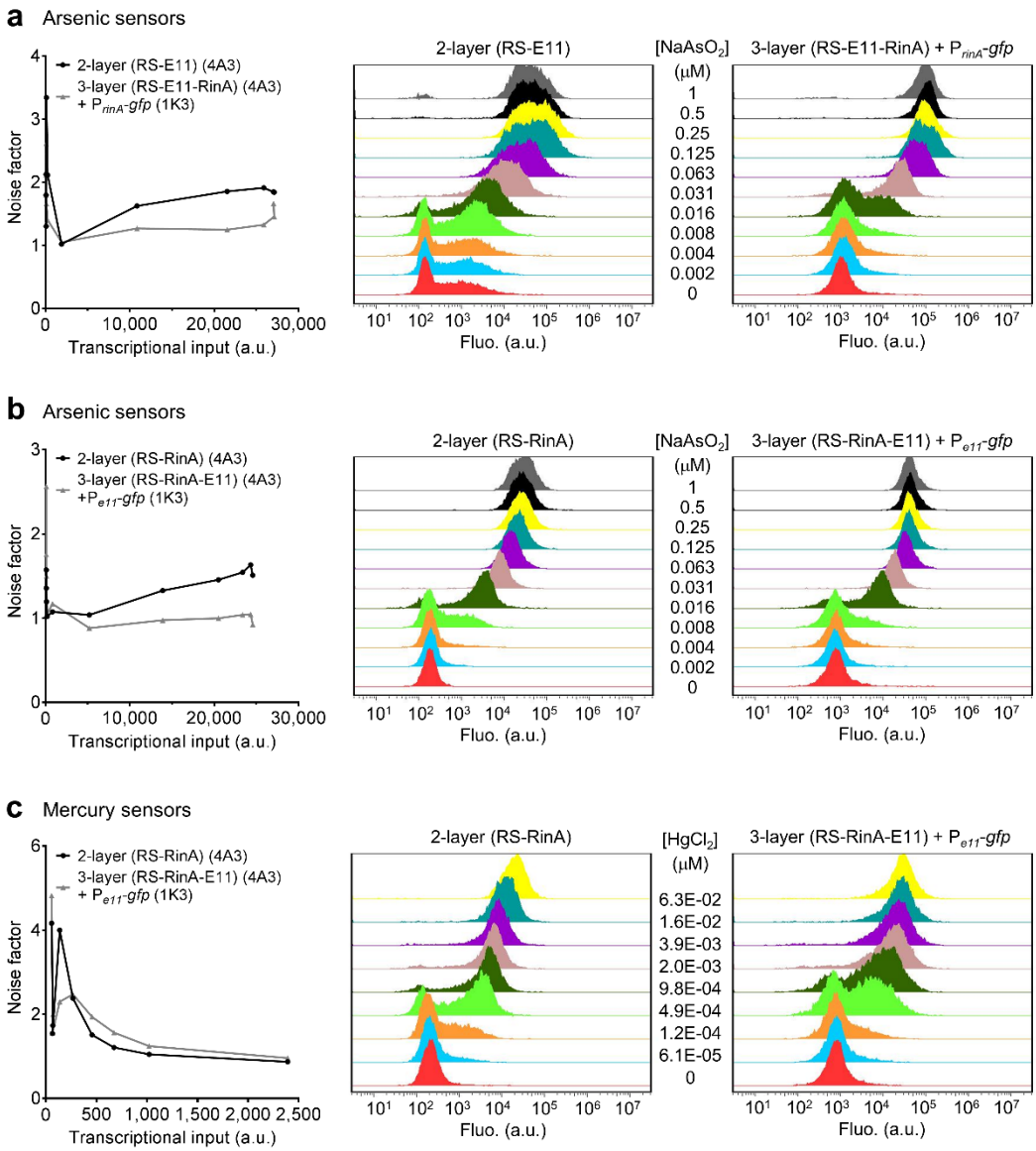


Figure 3.12: Noise characteristics of sensors with multi-layered amplifiers.

Related to **Figure 3.6b,c**, **Figure 3.7a** and **Figure 3.8c**. Noise characteristics and single cell flow cytometry assays of the arsenic (**a** and **b**) and mercury (**c**) sensors with multi-layered amplifiers. a.u., arbitrary units.

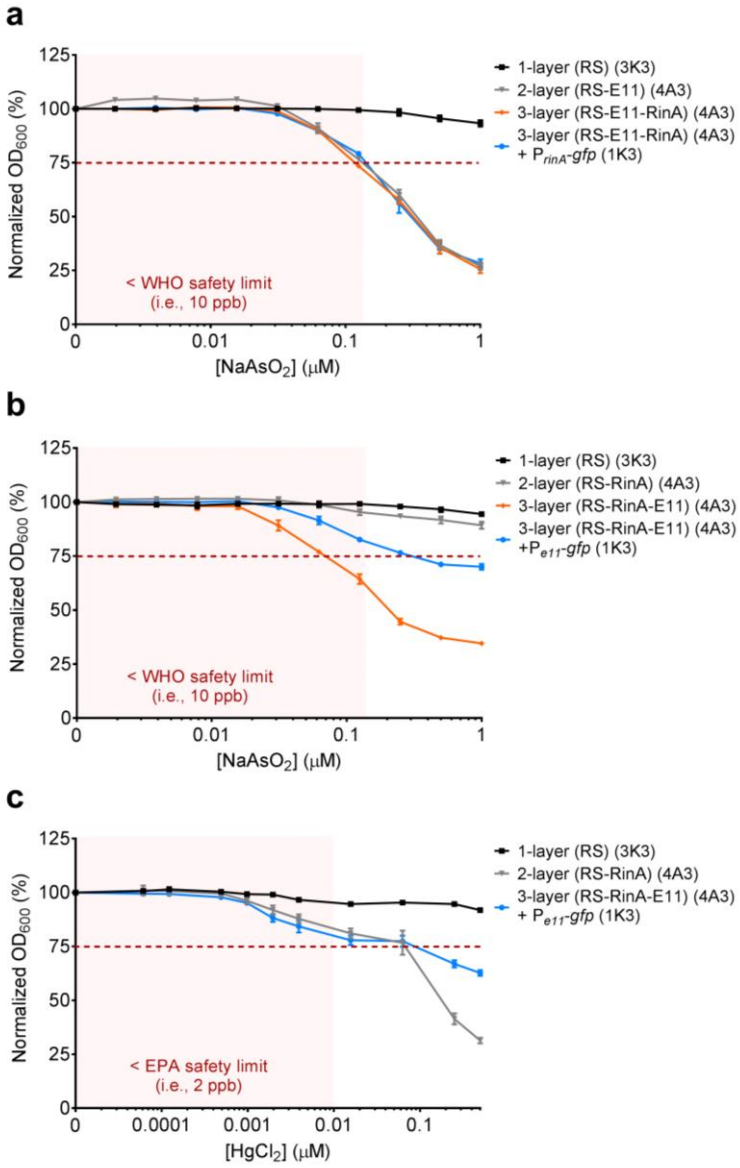


Figure 3.13: Load of engineered constructs on sensor hosts.

Related to **Figure 3.6b,c**, **Figure 3.7a** and **Figure 3.8c**. Normalised OD_{600} of arsenic (**a** and **b**) and mercury (**c**) sensors in response to various induction of arsenic or mercury, indicating the load of the engineered constructs imposed on the sensors. Normalised OD_{600} = (the OD_{600} of induced sample) / (the OD_{600} of uninduced sample). Red shaded regions indicate the input ranges that are lower than the WHO/EPA safety limit of the cognate pollutant. Error bars, s.d. (**a**, n = 4, **b** and **c**, n = 3).

It is noteworthy that the amplifiers not only increased the coupled sensor's output level, but also improved its detection limit. In general, a sensor's detection limit is largely determined by the input sensor module (Wang et al. 2015). In this study, single-layer $Amp30^E$ improved the sensor detection limit by 10 to 15-fold, while the multi-layered amplifier improved it

further. This could be due to the mode of action of transcription in nature. It is known that transcription events occur in bursts (Choi et al. 2008), which can be captured by the downstream ultrasensitive transcriptional amplifiers. Therefore, even at low level of input sensor induction, the transduced transcription bursts can still be amplified though at lower frequency, leading to detectable output reporter expression and thus lowering of the sensor detection limit. This assumption seems consistent with the two cell populations observed in the flow cytometry assay (**Fig. 3.12, 3.14 and 3.15**).

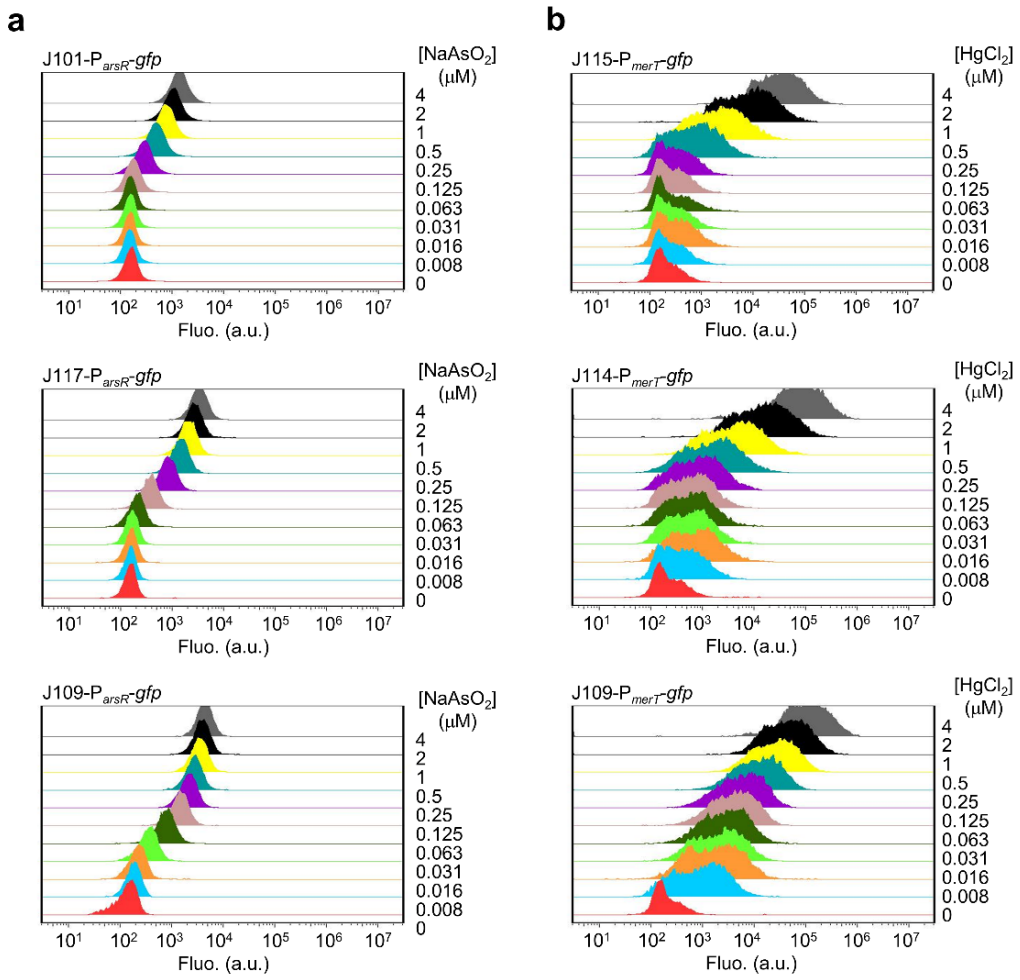


Figure 3.14: Sensor dose-response at single cell level under different receptor concentrations and ligand inductions.

Related to **Figure 3.2b** and **Figure 3.8a**. (**a**) Single cell flow cytometry data showing the cellular responses of the arsenic sensor induced by varied $NaAsO_2$. Strains shown containing J101-*arsR* (top), J117-*arsR* (mid), J109-*arsR* (bottom) construct respectively. (**b**) Single cell flow cytometry data showing cellular responses of the mercury sensor induced by varied $HgCl_2$. Strains shown containing J115-*merR* (top), J114-*merR* (mid), J109-*merR* (bottom) construct respectively. a.u., arbitrary units.

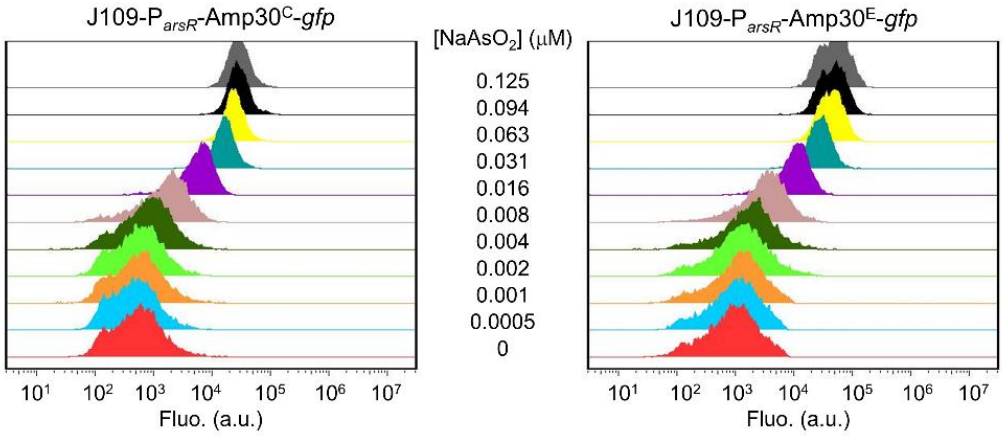


Figure 3.15: Single cell assay of the amplifier Amp30^{C} and its optimised version Amp30^{E}.

Related to **Figure 3.4**. Single cell flow cytometry data showing the cellular responses of the single-layer amplified arsenic sensors induced by varied $NaAsO_2$ concentrations.

In addition, I investigated two approaches to address the issue of sensor background leakage. The first approach is based on rational promoter engineering which can modulate the sensor's leakage level, sensitivity and dynamic ranges, and should be applicable to any ArsR-like repressor based sensors. The second approach utilises protease-based reporter degradation regulation to lower the sensor basal background without reducing the maximum output level and sensor sensitivity. Though more complex, this approach should be broadly applicable to different types of input sensors to regulate any protein level in the amplifier cascade. Other 'antagonists' can also be used in this platform to regulate protein expression, such as antisense RNAs, anti-sigma factors (Rhodius et al. 2013), and protein inhibitors (e.g., HrpV inhibiting HrpS) (Wang et al. 2014). By using orthogonal 'antagonists', it will be possible to regulate every transcription factor and output expression simultaneously with a single input, thus providing an alternative method for engineering and manipulating output expression.

Chapter 4. Designing a Cell-based Sensor Array for Field Testing

Although many biosensors have been demonstrated successfully in the laboratory, very few have reached the market. Besides biosafety concerns and unsatisfactory sensing performance in the natural environment, there is also a shortage of durable, inexpensive and user-friendly platforms for cellular biosensor storage and large-scale sample screening (Prindle et al. 2012; Volpetti et al. 2017). The majority of biosensor-based detection methods use a single biosensor to report its target with a single reporter. The sensor can report its target quantitatively by fitting its output level to a calibration curve (as for the sensors tested in this study); alternatively, the sensor can be designed to only report a small range of target concentrations (Rubens et al. 2016). However, the former always needs calibration as the cell number and cellular activity may vary under different conditions, while the latter requires serial dilution of samples. Easier read-outs such as comparing resulting output with standardised colour tables are still difficult and inaccurate.

Wackwitz et al. (2008) have proposed a 'traffic light' arsenic assay which could function independently of external calibration (van der Meer and Belkin 2010). The idea is to use a set of biosensors with different detection limits to generate 'traffic light' reporting patterns upon sensing their target (**Fig. 4.1a**). This assay was tested about ten years ago, and the set of biosensors was generated by tuning the RBS of their reporter gene (Wackwitz et al. 2008). However, this tuning method was not an efficient way to generate a large set of biosensors, and the tested sensor set could only detect a limited range of arsenic concentrations (e.g., can only distinguish among 0, 6, 10 and 100 ppb of arsenic using LacZ reporter, or 0, 4 and 50 ppb of arsenic using CCP reporter). Up to now, the 'traffic light' system has not been demonstrated successfully elsewhere, and no adequate method has been published to generate a large set of sensors for the system. Moreover, the assay was still based on bacterial liquid culture or cells growing on agar plates, which are not convenient for field testing.

Fortunately, the signal amplification approach along with the leakiness regulation developed in this study (Chapter 3) allowed the creation of numerous biosensors with various detection limits. Based on the sensor variants I generated here and hydrogel (Buffi et al. 2011; Shin 2012) or microfluidic entrapment (Volpetti et al. 2017), I designed and validated an innovative microbial sensor array that can generate graded volume-bar like patterns in response to varying

arsenic pollutant levels (**Fig. 4.1b**). Such easy-to-interpret patterns from the sensing array can be simply visualised by a cell phone camera without the need for sophisticated equipment, facilitating its use for portable, low-cost environmental monitoring in the field.

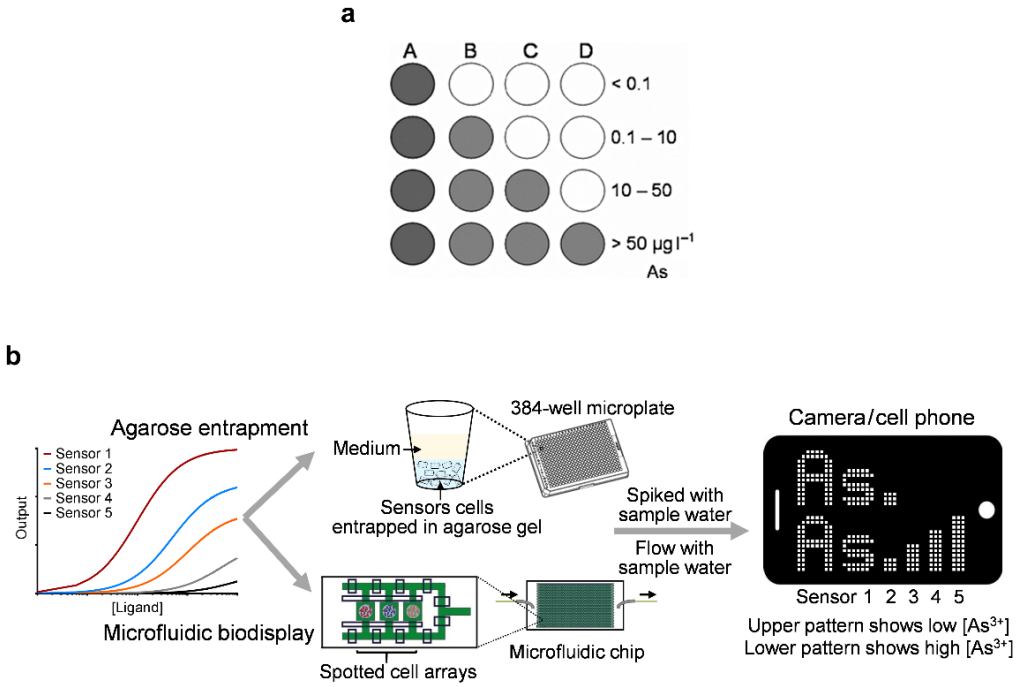


Figure 4.1: Diagrams of cell based sensor cell arrays.

(**a**) A 'traffic light' sensing system proposed previously. A–D represents arsenic sensors with sensitivity from high to low, grey circles represent sensors with output at ON state while the white circles represent sensors with output at OFF state (Wackwitz et al. 2008). (**b**) Sensors (i.e., Sensor 1–5) with varying sensitivities and outputs can be generated by the 3-step signal amplification along with leakiness regulation (see Chapter 3). Utilising agarose entrapment or microfluidic encapsulation, they are used to build a microbial sensor array which displays an easy-to-read volume bar-like pattern for mobile phone-based easy-to-use and accurate field monitoring of target environmental contaminants such as arsenic in drinking water.

4.1 Designing and testing sensor array for arsenic detection

Several arsenic sensors were compared in liquid culture, and eight of them with a range of detection limits (**Fig. 4.2a,b**, **Appendix Table 2.2**) were selected to test their performance under agarose gel entrapment (see Section 2.5.2, **Fig. 4.2c–e**). The agarose gel-entrapped sensors were spotted and induced with $NaAsO_2$ in a 384-well microplate as shown in **Figure 4.2c**, and were measured by fluorometry to quantify the fluorescent output (**Fig. 4.2d**). A photograph was taken using a Gel Doc imaging system to visualise the output (**Fig. 4.2e**). The fluorometry measurement shows that the gel entrapped sensors exhibited similar relative

sensitivities as those in liquid culture. However, due to the lower sensitivity of the gel doc imaging camera, some sensors (e.g., As4) appeared less sensitive in the image taken by the camera.

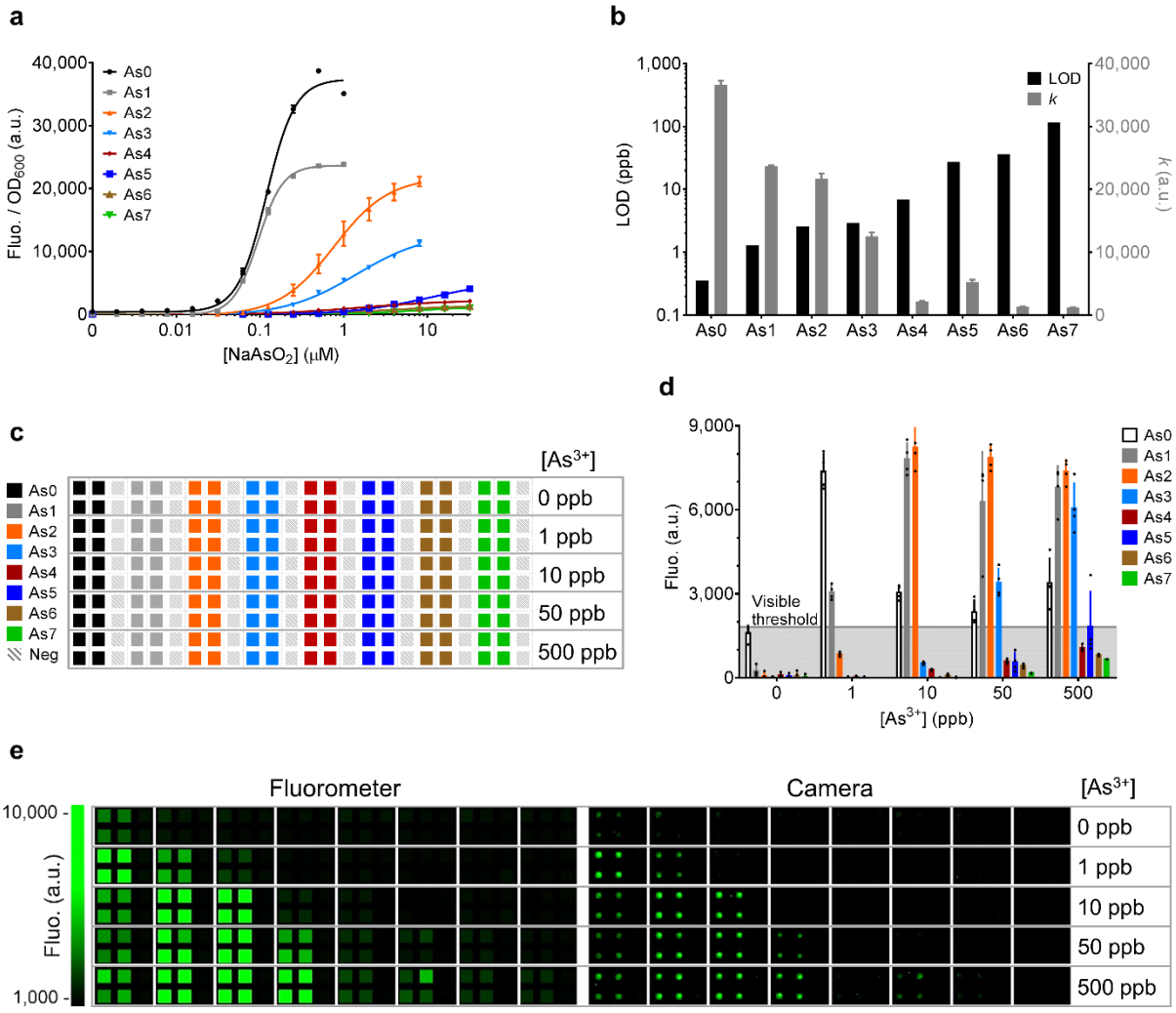


Figure 4.2: Characterisation of selected arsenic sensors entrapped in agarose gel.

(**a**) Characterisation of eight arsenic sensors of varying sensitivities in liquid culture. Detailed configuration of each sensor is described in **Appendix Table 2.2**. Error bars, s.d. (n = 3). (**b**) Limit of detection (LOD) and maximum output expression (*k*) of the eight arsenic sensors in **a**. Error bars, 95% confidence intervals. (**c**) Microplate well layout for testing eight arsenic sensors entrapped in agarose gel. Results are shown in **d** and **e**. (**d**) Quantified fluorometry data of the arsenic sensors in **c**. Data were collected after 24 h incubation at 37˚C, and calculated after subtraction of the fluorescent readouts of negative controls. By visually inspecting the agarose gel-entrapped sensors in **e** and comparing this with the fluorometry data, we determined an effective visible threshold to be 1,800 fluorescence units. Error bars, s.d. (n = 4, technical replicates). (**e**) Graph showing the response and the visibility of the agarose gel-entrapped sensors under various arsenic induction levels. Data were collected after 24 h incubation

at 37˚C. Left panel: fluorometry data. Right panel: photos taken by the Bio-Rad Gel Doc XR+ system. a.u., arbitrary units.

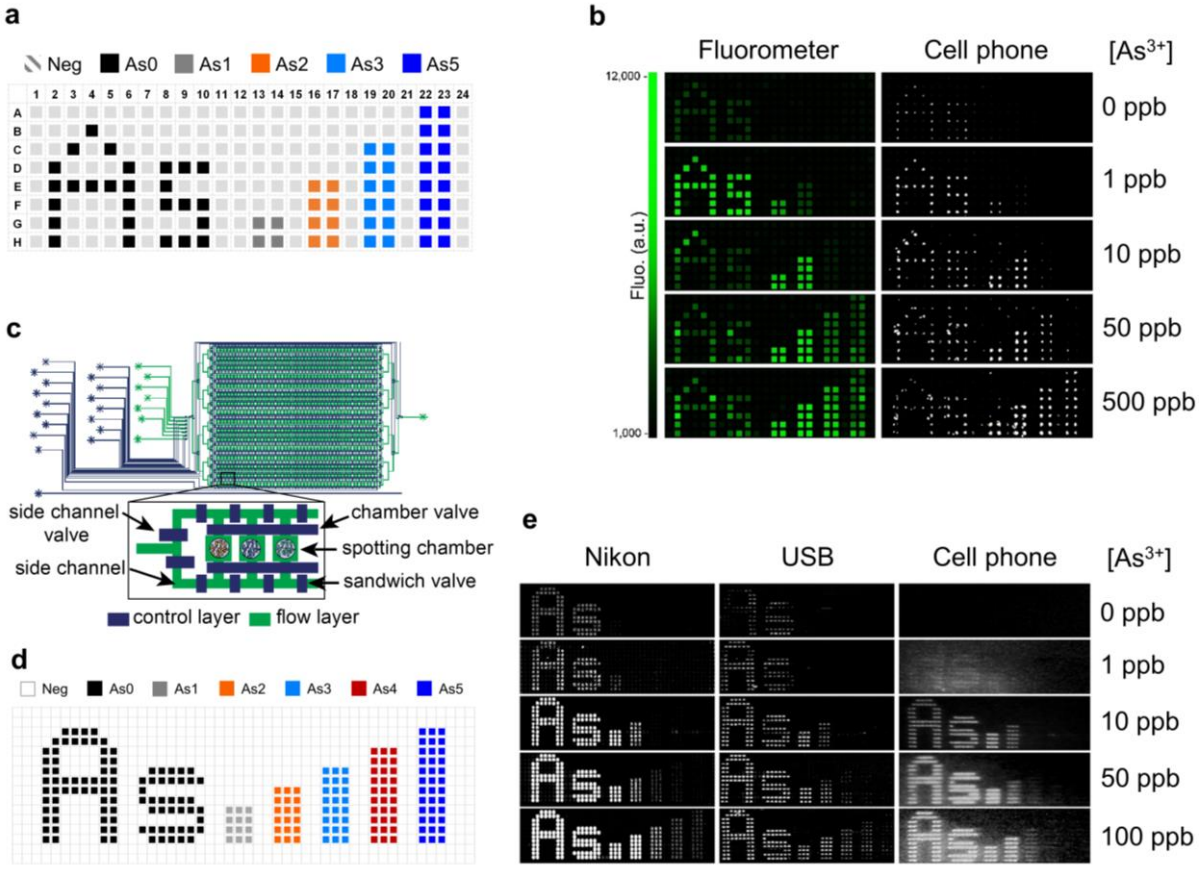


Figure 4.3: Microbial sensor array display enabled by agarose hydrogel entrapment and microfluidic encapsulation for easy-to-use monitoring of arsenic contamination.

(**a**) Design of the sensor array in a half 384-well microplate. 'As' symbol represents arsenic for identifying the type of contamination, and the number of volume bars can indicate the relevant arsenic concentration. (**b**) Agarose gel encapsulation-enabled microbial sensor array for monitoring arsenic contamination, showing different output response patterns upon various arsenic induction levels 24 h post incubation. Left panel: fluorometry data. Right panel: images taken by a cell phone. a.u., arbitrary units. (**c**) Schematic showing the design of a microfluidic device with spotted sensor cells in multiple independently controlled compartments. (**d**) Design of the arsenic sensor array based on microfluidic biodisplay. (**e**) Microfluidic encapsulation-enabled sensor array display for monitoring different levels of arsenic. Left panel: images acquired by a Nikon microscope. Middle panel: images acquired by a FITC USB fluorescence microscope. Right panel: images taken by a cell phone camera.

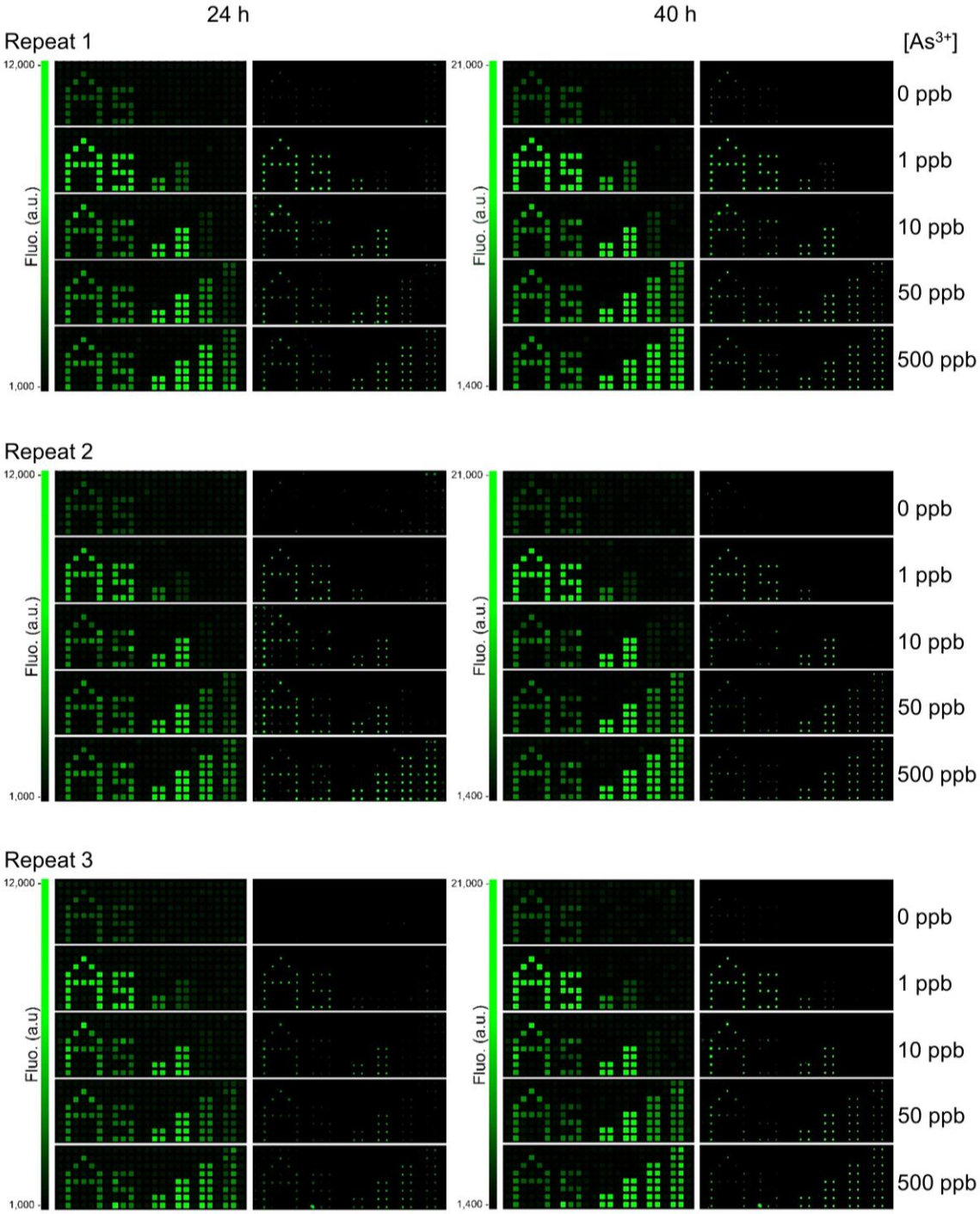


Figure 4.4: Three individual repeats of agarose gel-enabled microbial sensor array for detecting arsenic.

Three individual repeats showing the stability of the arsenic sensors entrapped in agarose gel following the sensor array designed in **Fig. 4.3a**. Data were collected from fluorometric measurement or a Gel Doc camera after 24 h and 40 h induction at 37˚C, showing increased fluorescence signals after longer incubation. a.u., arbitrary units.

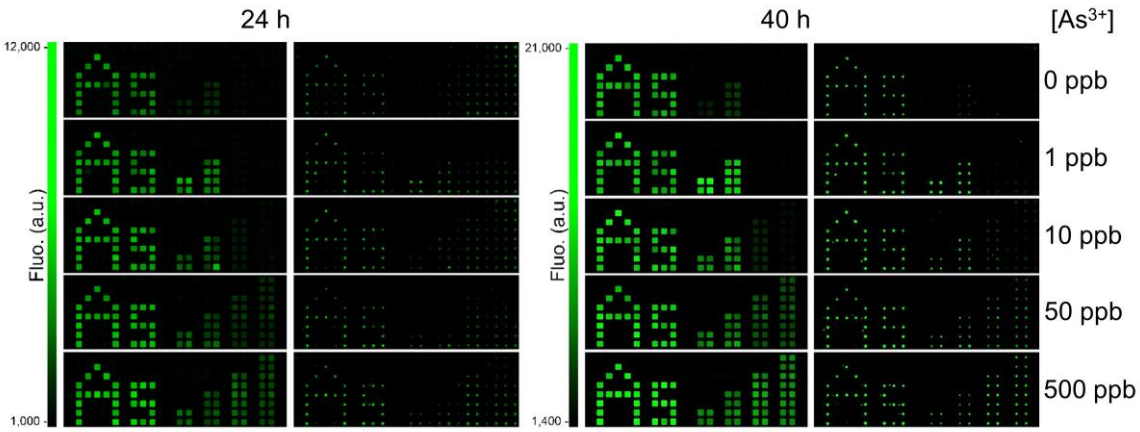


Figure 4.5: Test of the arsenic sensor array at room temperature.

Agarose gel-enabled microbial sensor array for detecting arsenic at room temperature (i.e., 25˚C). The array design is the same as in **Fig. 4.3a**. Data were collected from fluorometric measurement or a Gel Doc camera after 24 h and 40 h induction. a.u., arbitrary units.

I next designed and tested the sensor array entrapped in agarose gel using a subset of the eight characterised arsenic sensors (As0–3 and As5). They were spotted in a 384-well microplate following a layout designed to display a volume bar-like pattern (**Fig. 4.3a**). I tested the sensor array under various arsenic ($NaAsO_2$) induction levels after 24 h incubation. A cell phone camera was used to simplify the imaging procedure in addition to the fluorometric assay (see Materials and Methods). The expected volume bar-like fluorescent patterns were observed from both cell phone image and the fluorometric measurement, i.e.: no arsenic – no pattern, 1 ppb – 'As' with 1 bar, 10 ppb – 'As' with 2 bars, 50 ppb – 'As' with 3 bars, 500 ppb – 'As' with 4 bars (**Fig. 4.3b**). Three further repeats were performed on different days with similar results obtained, suggesting good stability of the agarose gel-based sensor array (**Fig. 4.4**). Additionally, I observed stronger fluorescent signals after longer incubation time (**Fig. 4.4**), and the agarose gel-entrapped sensors also worked at room temperature (**Fig. 4.5**). However, different testing conditions may generate different patterns, thus the sensor array should be re-arranged accordingly.

The sensor array was also tested in a recently developed microfluidic biodisplay (Volpetti et al. 2017) (see Materials and Methods, **Fig. 4.3c,d**). The microfluidic device contains 768 individually programmable biopixels, and each biopixel contains a chamber where different sensor cells can be spotted. Each chamber has valves around it to control the flow of media or samples to the sensors, and also to entrap the sensor cells. Francesca Volpetti, Ekaterina Petrova and Sebastian J. Maerkl from Ecole Polytechnique Federale de Lausanne designed

and performed the microfluidics-based experiments. Similar results of easy-to-interpret volume bar-like pattern were obtained from a Nikon fluorescence microscope and portable devices including both a low-cost USB microscope and a cell phone camera (**Fig. 4.3e**, **Appendix Fig. 4.1**).

4.2 Using the sensor array to test groundwater samples

To verify whether the agarose-based sensor array is functional in natural potable water, I proceeded to test it using groundwater samples that I collected from Khulna, Bangladesh (see Section 2.6).

4.2.1 Groundwater test using M9 medium

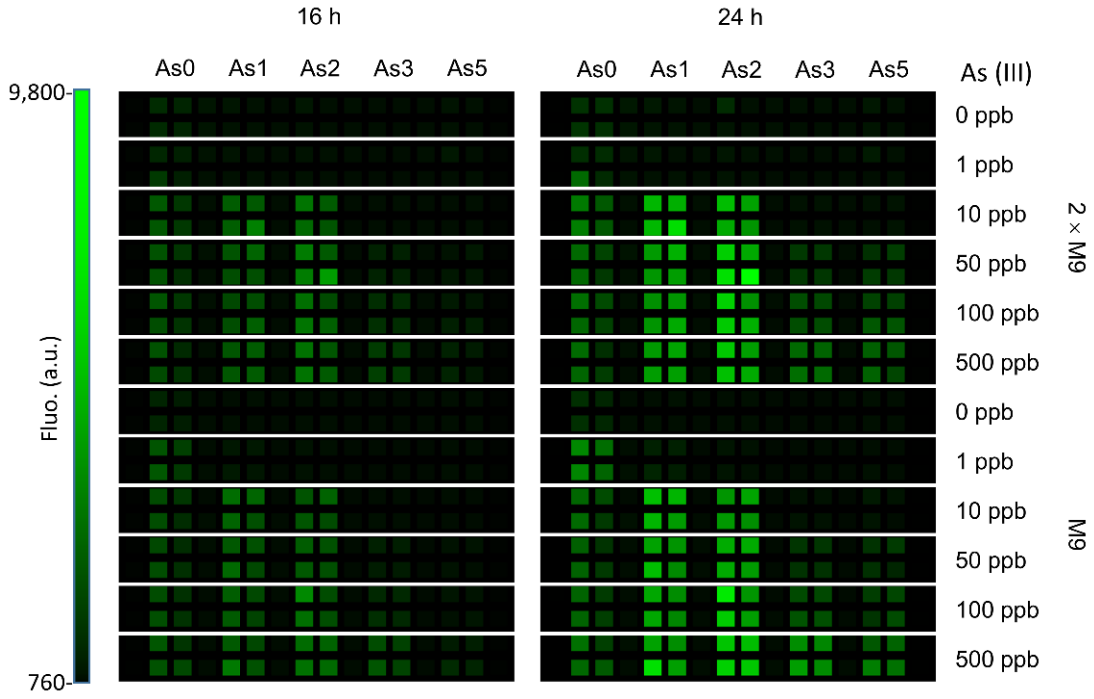


Figure 4.6: Comparing sensor response with different media preparation.

Comparing the response of agarose gel-entrapped arsenic sensors incubated in M9 or 2 × M9 medium diluted twofold with samples, showing that the different media preparation did not affect the sensor response. Fluorometry data were collected after 24 h and 40 h induction at 37˚C. a.u., arbitrary units.

As the water sample will dilute the culture medium, I concentrated the medium 2-fold, and mixed the water sample with concentrated medium with a ratio of 1 : 1. By doing this, although the water sample will be diluted two-fold, the final concentration of the medium will be the

same as in previous tests. I first confirmed that the concentrated medium after dilution worked as same as the previous one (**Fig. 4.6**).

I next quantified the arsenic levels in 22 groundwater samples using ICP-MS, and tested 4 of them using the sensor array (see Section 2.6, **Fig. 4.7b**). 50 ppb arsenic ($NaAsO_2$) induction was performed as a positive control. The sensor array showed a slightly stronger 'As' pattern with Sample 5 (9 ppb arsenic after dilution) than the non-induced sensor array, and more obvious patterns with Sample 11 (50 ppb arsenic after dilution) and Sample 21 (85 ppb arsenic after dilution). These suggest that the sensor array can at least respond to water samples with arsenic ≥ 18 ppb, which is lower than the WHO's guideline for the drinking water in Bangladesh (i.e., 50 ppb). However, the sensitivity was not as good as the laboratory test using ddH_2O spiked with $NaAsO_2$.

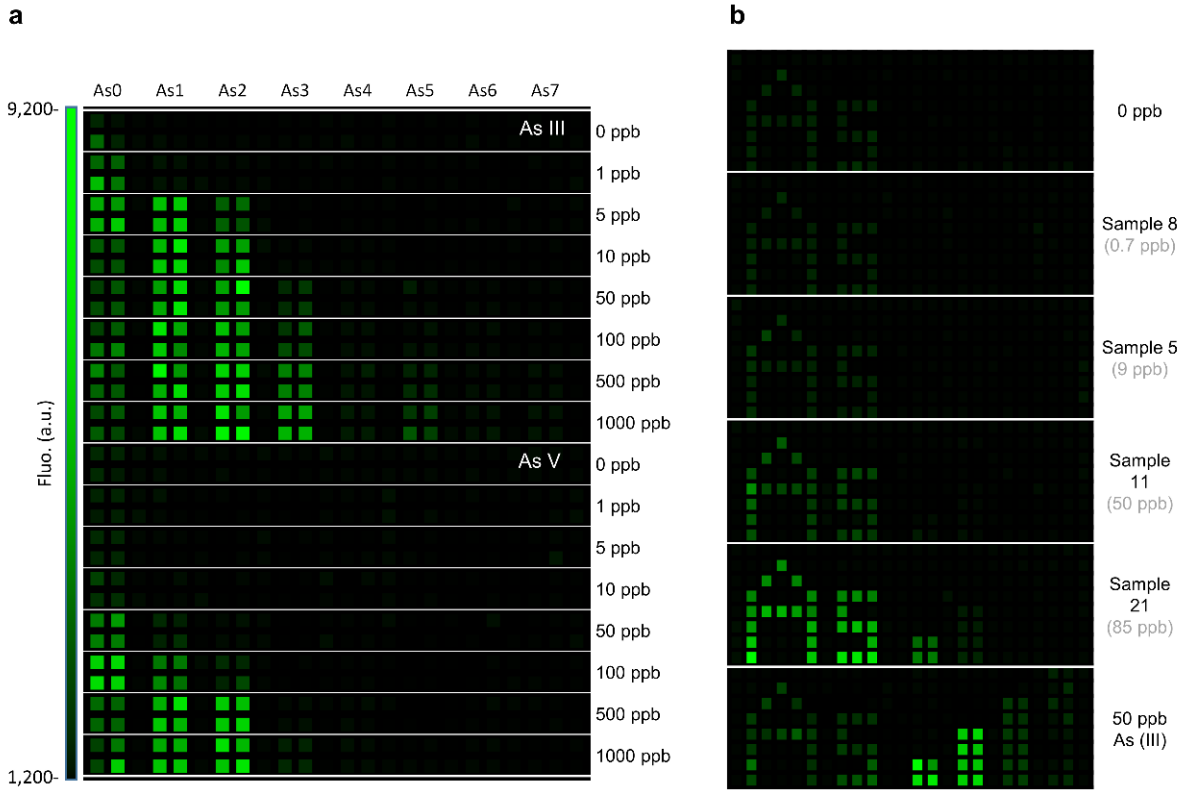


Figure 4.7: Test of groundwater samples using agarose gel-based arsenic sensor cell array.

(**a**) Comparing the response of agarose gel-entrapped arsenic sensors to arsenite ($NaAsO_2$) and arsenate (Na_2HAsO_4) induction. (**b**) Testing groundwater samples using agarose gel-based arsenic sensor cell array designed in **Figure 4.3a**. Test of 50 ppb arsenic ($NaAsO_2$) was included as a positive control. Fluorometry data shown were collected after 24 h incubation at 37˚C. a.u., arbitrary units.

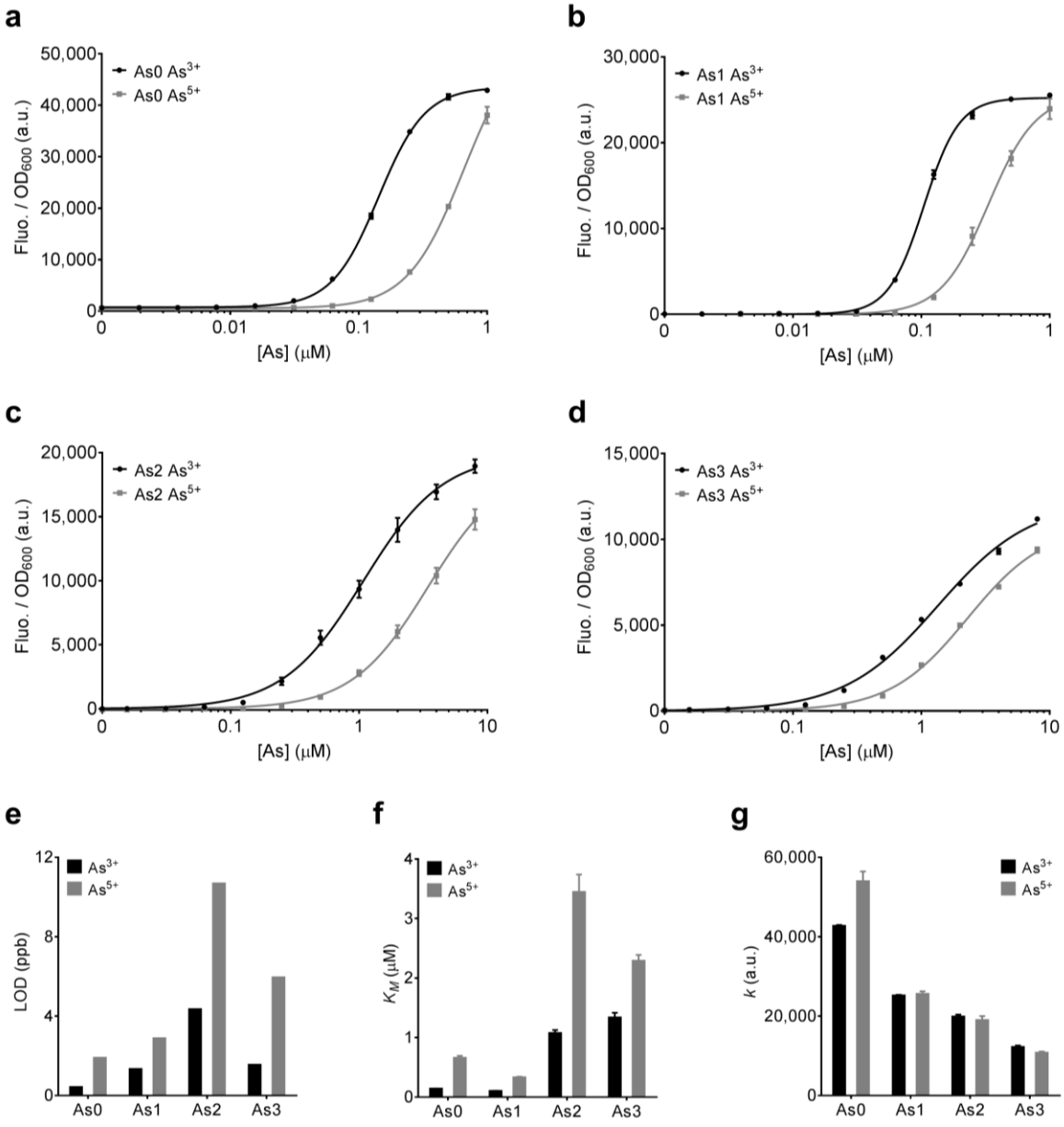


Figure 4.8: Comparing the dose-response of arsenic sensors in the presence of arsenite and arsenate.

(**a**–**d**) Characterisation of arsenic sensors As0–3 in response to the same concertation of arsenite ($NaAsO_2$) or arsenate (Na_2HAsO_4). Error bars, s.d. (n = 3). (**e**–**g**) Limit of detection (LOD), Hill constant (K_M) and the maximum output (k) of the selected arsenic sensors in response to arsenite and arsenate. Error bars, 95% confidential intervals. a.u., arbitrary units.

There might be two reasons that caused the sensitivity decrease. One might be that the extra elements in natural water inhibited the sensors' performance. The other might be that the tested water samples mainly contained arsenate; the sensors might sense the arsenate with less sensitivity than arsenite. To identify which one was the real reason, I compared the sensor response to ddH_2O spiked with arsenate (Na_2HAsO_4) and arsenite ($NaAsO_2$) in both liquid culture (**Fig. 4.8**) and in agarose gel (**Fig. 4.7a**). Indeed, the sensors were less sensitive to arsenate than to arsenite as described previously (Stocker et al. 2003). Moreover, the sensor array induced with arsenate showed similar patterns to the ones induced with the water sample at similar arsenic concentrations (**Fig. 4.7a**). This suggests that the tested water samples were likely containing arsenate only and did not cause significant inhibition on the sensors. To further confirm this, the arsenate and arsenite can be quantified using X-ray Photoelectron Spectroscopy (Chowdhury et al. 2011).

4.2.2 Medium optimisation for groundwater test

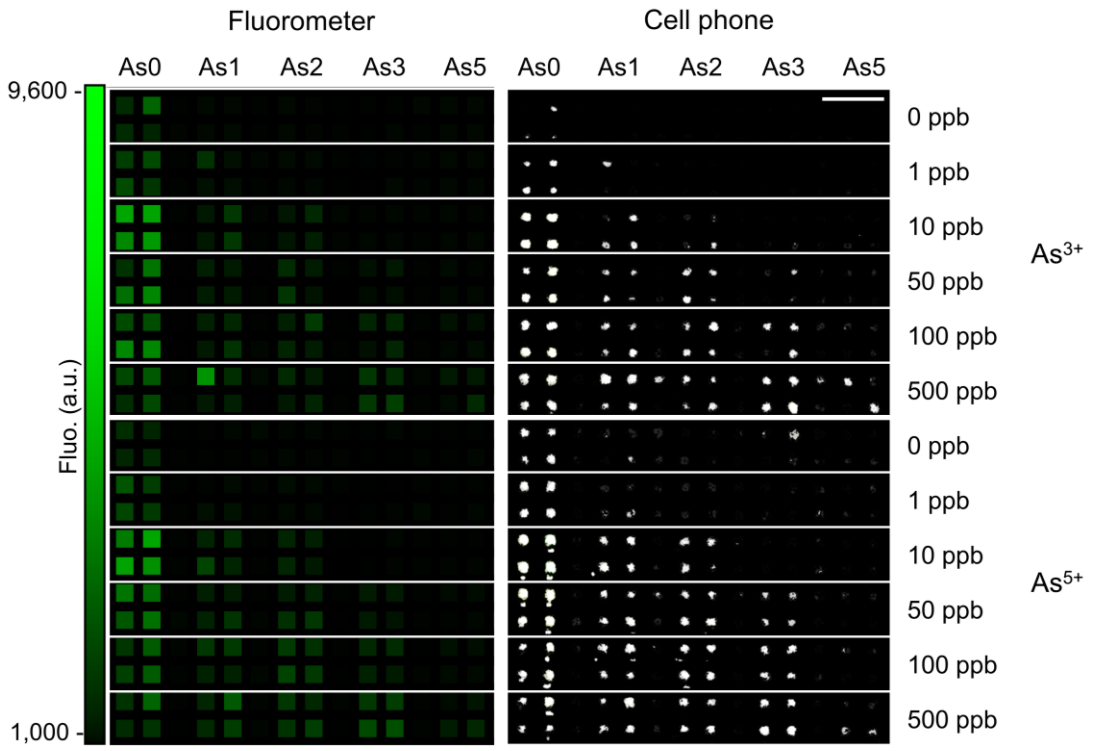


Figure 4.9: Comparison of arsenic sensors' responses to arsenite and arsenate in an optimised MOPS medium.

Characterisation of arsenic sensors As0–3 and As5 in response to the same concentration of arsenite ($NaAsO_2$) or arsenate (Na_2HAsO_4) in the optimised MOPS rich medium (with 1/4 PO_4^{3-} and 0.2% (v/v) glycerol). Scale bar, 1 cm. a.u., arbitrary units.

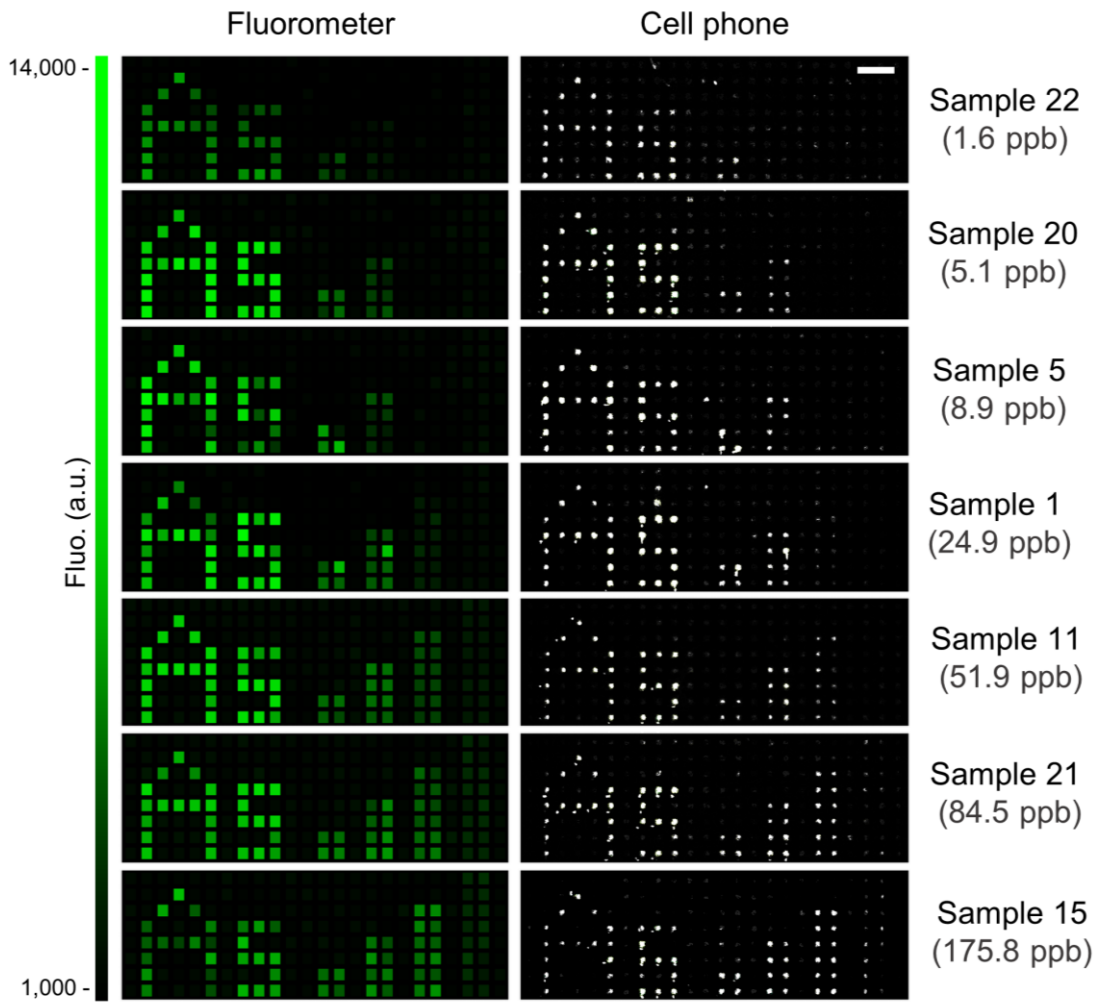


Figure 4.10: Agarose gel encapsulation-enabled microbial sensor array for monitoring arsenic contamination from groundwater samples.

The water samples were mixed with an optimised MOPS medium (with 1/4 PO_4^{3-}) at 1 : 1 ratio (see Section 2.5.2 and Section 2.6). The arsenic concentrations after two times dilution are shown in brackets. Scale bar, 1 cm. a.u., arbitrary units.

Previous results suggested that the groundwater samples from Bangladesh mainly contained arsenate, to which the arsenic sensors were less sensitive (**Fig. 4.7** and **4.8**). Comparing the data from LB-based liquid culture tests (**Fig. 4.8**) and M9-based agarose gel tests (**Fig. 4.7**), the sensors' sensitivities to arsenate and arsenite were not in the same proportion: the agarose-entrapped sensors were much less sensitive to arsenate. This means the arsenic sensors' responses to arsenate were further inhibited by another factor. As previously used M9 medium contains high level of phosphates, which share the same uptake systems as arsenate (Willsky and Malamy 1980), it maybe the factor that reduced the sensors' sensitivities to arsenate.

To validate above hypothesis and to improve sensors' performance in agarose gel, optimised M9 and other media (i.e., LB and MOPS) have been tested (see Section 2.5.2, **Appendix Fig. 4.2**). The results showed that both LB and MOPS media, as well as M9 medium with lowered phosphates improved the sensors' sensitivities to arsenate, confirming that the high level of phosphates was the limiting factor. Moreover, the arsenic sensors had similar sensitivities to arsenate and arsenite in both LB and MOPS media. Given the low background and high output upon arsenic induction, an optimised MOPS (see Section 2.5.2 and Section 2.6) was selected for the groundwater samples test. This medium maximised the sensors' responses to arsenate (**Fig. 4.9**), and allowed the sensor cell array to robustly test real environmental samples with arsenic ≤ 3.18 ppb (**Fig. 4.10**). The comparison of ICP-MS with arsenic sensor cell array-determined arsenic concentrations for the water samples are listed in **Table 4.1**.

Table 4.1: Comparison of ICP-MS with arsenic sensor cell array-determined arsenic concentrations for Bangladesh groundwater samples

	Arsenic measured by ICP-MS		Diluted concentration used for arsenic sensor cell array	Arsenic measured by arsenic sensor cell array
Sample	ppb	% r.s.d	(ppb)	(ppb)
1	49.84	7.27	24.92	10 − 50
5	17.82	4.79	8.91	1 − 10
11	103.70	1.36	51.85	50 − 100
15	351.60	0.14	175.80	100 − 500
20	10.20	12.38	5.10	1 − 10
21	169.00	1.96	84.50	50 − 500
22	3.18	14.52	1.59	1 − 10

r.s.d: relative standard deviation.

4.3 Discussion

To facilitate the application of cell-based sensors, I developed a sensing platform using the engineered sensor variants based on agarose gel entrapment. This platform can display an easy-to-interpret volume bar-like pattern to indicate the cognate pollutant level in the samples. I demonstrated that the sensor array's output patterns can be simply captured by a cell phone camera without the need for sophisticated equipment, facilitating its potential use as a portable, low-cost environmental monitoring tool in the field. The present detection range of the sensor array used is 1 – 500 ppb arsenite, but this can be expanded or refined by varying the number and types of the arsenic sensor variants used to create the cell arrays. More strikingly, though with lower sensitivity, it can also detect ddH_2O spiked with ≤ 1 ppb arsenate and groundwater samples containing ≤ 3.18 ppb arsenic. Furthermore, the sensor cell array can distinguish true negative responses from false negative responses caused by lethal dose of antimicrobials (See Section 2.5.2, **Fig. 4.11**).

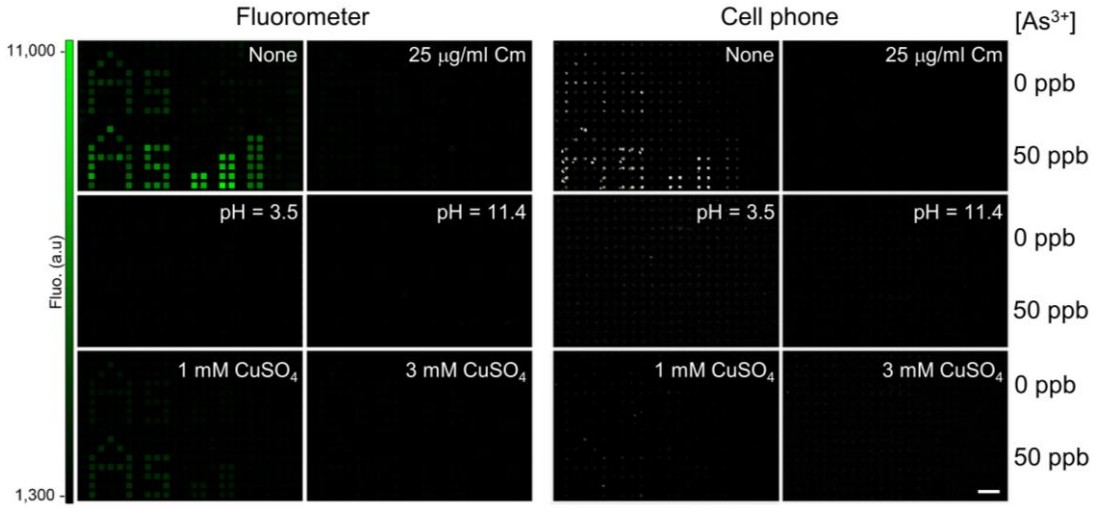


Figure 4.11: Characterisation of agarose gel entrapment-enabled microbial sensor array under lethal antimicrobial conditions.

Cells were incubated in M9 medium with 25 μg/ml chloramphenicol (Cm), under low pH (pH = 3.5), high pH (pH = 11.4), with 1 mM or 3 mM $CuSO_4$. The sensor arrays were induced with 0 or 50 ppb As^{3+} of $NaAsO_2$. Scale bar, 1 cm. a.u., arbitrary units.

Mercury sensors have also been tested in agarose gel-based and microfluidic-based sensing platforms, but the currently selected sensors were not sensitive enough for real applications (data not shown). Better sensors need to be identified as mentioned above, and other relevant methods may also be applied to improve the sensing performance.

In contrast to typical sensors having a single colorimetric output, the sensor cell array designed here makes it easier and quicker for end users to tell the type and level of the cognate contaminant in the sample, and could be readily adapted to other cellular biosensors. Though the current sensor cell array requires 24 h of incubation to generate sufficient fluorescence output for visualisation, further optimisation can be done to reduce the processing time if needed. For example, cell densities can be optimised to increase fluorescence levels so that the output can reach the visual detection threshold rapidly, and fast enzymatic reaction-based outputs (e.g., NanoLuc (Cevenini et al. 2018) or LacZα peptide (Ma et al. 2018)) can speed up the detection processes. Importantly, by entrapping sensor cells inside hydrogels or microfluidic devices, it significantly reduces the chance of their escape to the open environment and helps mitigate the biosafety concerns for their final field application. With the new advances in cell entrapping materials and preserving methods, the engineered microbial sensor array has the potential to be stably stored for long term use (Buffi et al. 2011; Volpetti et al. 2017; Liu et al. 2018b).

Chapter 5. Developing Biosensors in Cell-free Expression System

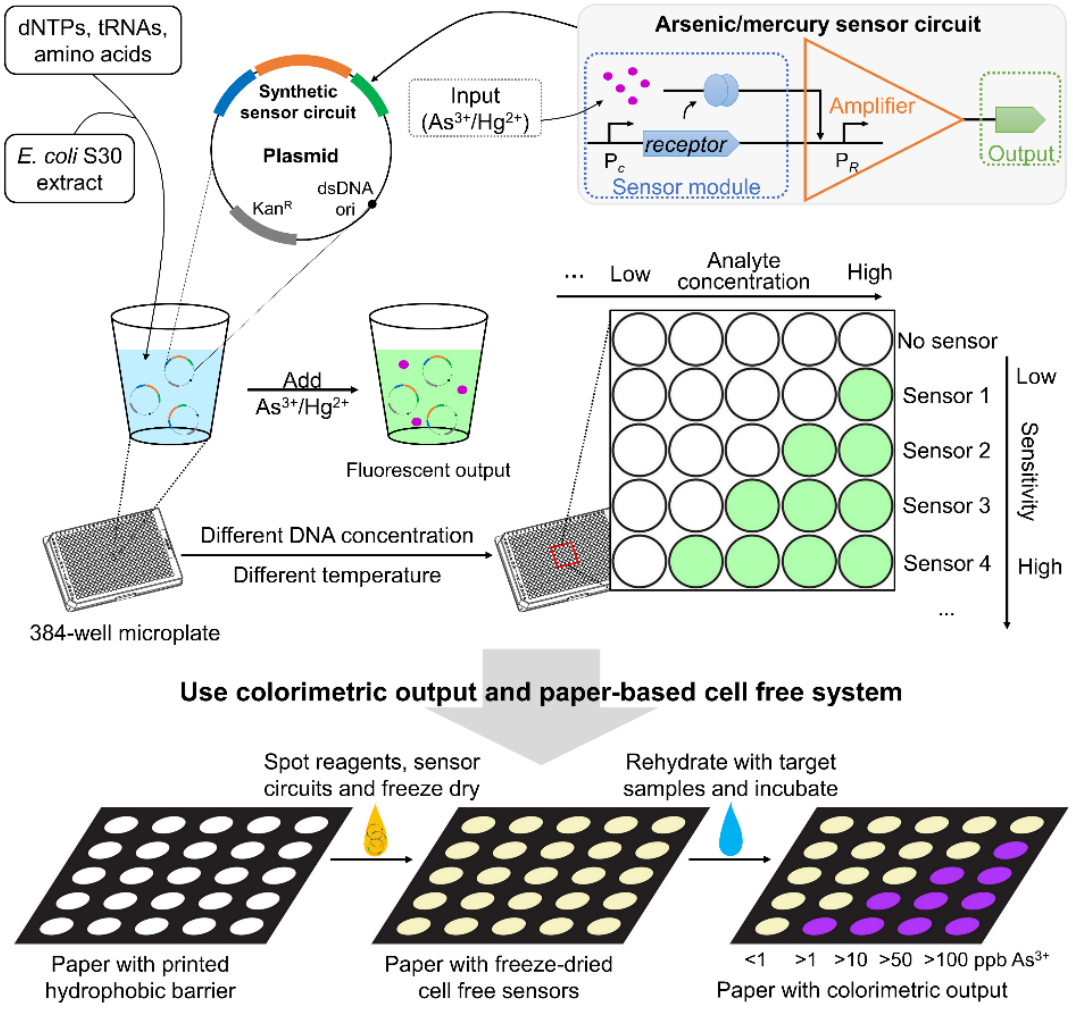


Figure 5.1: Experimental process for testing arsenic and mercury sensors in CFS.

The aim of the CFS-based sensor project is to make a sensor array on paper which can be used to detect arsenic/mercury and also to determine their concentrations in drinking water. To achieve this, a set of arsenic/mercury sensors will be tested and optimised in cell-free solutions using 384-well microplates, but with fluorescent output for easy measurement. Then the working sensor sets will be tested on a paper-based CFS, using colorimetric output for easy visualisation. P_C, constitutive promoter, P_R, cognate promoter of the receptor.

Cell-free transcription-translation (TX-TL) system is becoming a favourable technology for *in vitro* synthetic biology study due to its advantages of flexibility, fast prototyping as well as constructing minimal cells (Chappell et al. 2013; Niederholtmeyer et al. 2015; Garamella et al. 2016; Karim and Jewett 2016; Perez et al. 2016). Adapting cell-based biosensors to such cell-free system (CFS) is also becoming a popular way in terms of minimising biosafety concerns for sensor application (Pellinen et al. 2004; Pardee et al. 2014; Pardee et al. 2016a; Didovyk et al. 2017). Compared to cell-based biosensors, the CFS-based sensors can be more tolerant to toxic translational products (Perez et al. 2016) and more sensitive, and respond faster to their targets. Also, the recently developed portable paper-based cell-free sensing platform can be affordable and easy-to-use for many biosensor applications (Pardee et al. 2014; Pardee et al. 2016b). Due to these advantages, I transferred the arsenic and mercury sensors from cells to CFS, and aimed to develop paper-based sensors for detecting arsenic and mercury water contaminants.

The procedure for developing CFS-based biosensors is described in **Figure 5.1**. Basically, the arsenic and mercury sensors were first tested and optimised in a solution-based CFS, then transferred to a paper-based CFS with colorimetric output. Further, I aimed to develop sensors with different detection limits in CFS, which can be used in a 'traffic light' sensing system. In contrast to colour-based test kits (e.g., EZ arsenic test kit) and typical biosensors having a single colorimetric output, the 'traffic light' sensor array will make it easier and more accurate for end users to tell the levels of contaminants in water samples.

5.1 CFS-based arsenic biosensors

5.1.1 Failed initial test of CFS-based arsenic sensors

I first tested previously studied arsenic sensors in cell extract-based CFS (CE-CFS). Those sensors were from the first step optimisation (Chapter 3.1 and Wang et al. 2015), where ArsR expression level was regulated by different constitutive promoters (i.e., J101/105/114/117) and *gfp* was used as reporter. A positive control was made to determine the maximum level of P_{arsR} expression in CFS, and also to evaluate the activity of the CFS in each experiment.

The arsenic sensors were induced with 1 μM $NaAsO_2$ (i.e., induced samples) or nuclease free water (i.e., non-induced samples) in CFS. Unfortunately, for each sensor, the *gfp* expression level of induced samples were not distinguishable from the one of non-induced samples (data not shown). Even the difference between the negative control and the positive control was not significant (data not shown). This may because P_{arsR} is a weak promoter; also, CE-CFS has strong background green fluorescence, which may have obscured the weak *gfp* expression.

5.1.2 Changing reporter and regulating ArsR : P_{arsR} ratio

To minimise the interference from the background fluorescence in CFS, I replaced *gfp* with *mCherry*. Furthermore, to simplify tuning of sensor detection limit and basal expression (leakage), I separated the sensor module and the output module, and moved them to two plasmids (**Fig. 5.2a**). I tested the sensor with different receptor : promoter ratios (**Fig. 5.2b–e**). No background red fluorescence was observed in the CFS. For all the four receptor : promoter conditions, the sensor responded slightly to arsenic induction; however, the basal expression was always high even with the ratio > 8. Similar experiments have been repeated with other receptor circuits (i.e., *arsR* expression driven by other promoters), and similar results were obtained (data not shown).

The high basal expression suggests that P_{arsR} cannot be repressed sufficiently by ArsR in CFS under the current testing conditions. To improve ArsR-P_{arsR} binding efficiency, I tested following strategies:

1) It has been shown that proteins tend to bind to DNA more tightly at low temperature due to thermodynamic effects (Hauschild et al. 2005; Merrin et al. 2011). Therefore, I incubated ArsR with P_{arsR} at low temperature to improve their binding.
2) As an extra ABS can reduce the basal level of the arsenic promoter in cells (Section 3.4), this was also tested in CFS.
3) One significant difference between cellular and CFS environments is the macromolecular crowding conditions. As the cell membranes limit the cell volumes, macromolecules are highly crowded which may not be the case in CFS. Therefore, to enhance the crowding effect in CFS, I added extra crowding agents (PEG and Ficoll) to CFS.
4) In P_{arsR}, ABS overlaps with a strong −35 region, which may affect the binding of ArsR to ABS. Therefore, weakening such σ^{70} binding site may improve ArsR-P_{arsR} binding affinity thus reducing the basal expression. Moreover, tuning the binding affinity of σ^{70} to a promoter regulates the output dynamic range (Chen et al. 2018), which may further improve sensor response.

Details of the experimental designs and results are shown in the following sections.

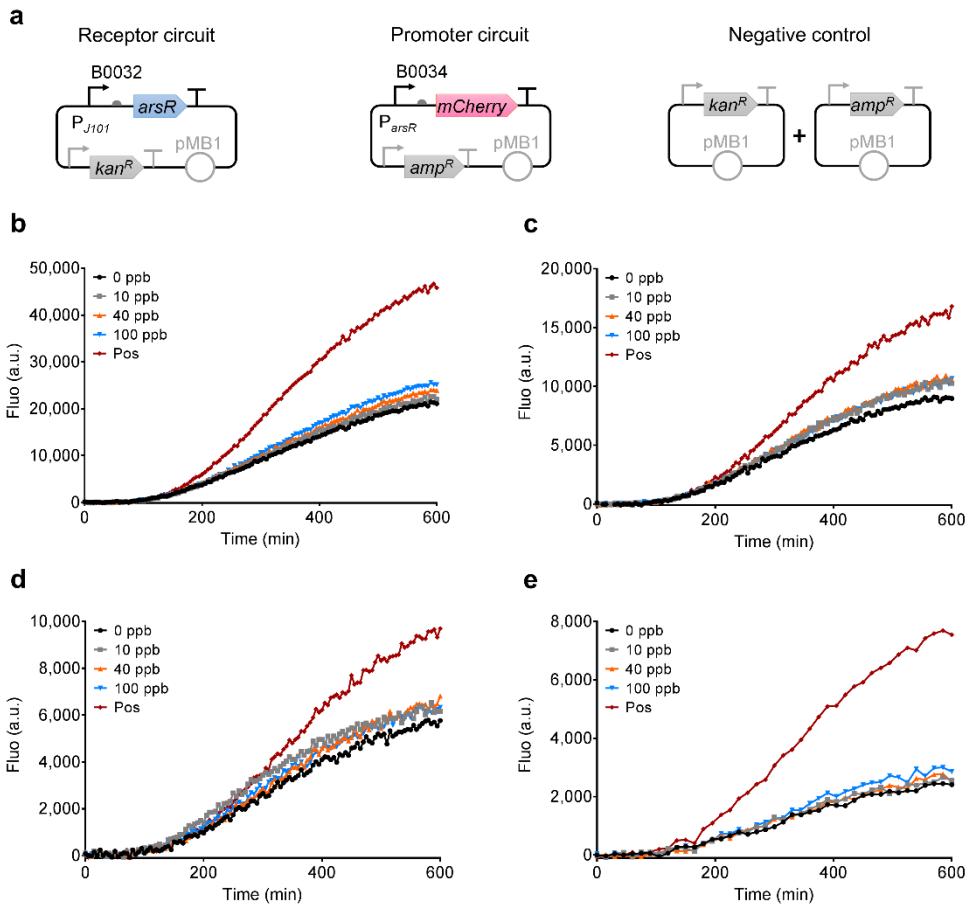


Figure 5.2: Tuning arsenic sensor background leakage in CFS by adjusting receptor : promoter ratio.

(**a**) Diagrams of genetic circuits used in this experiment. (**b**–**e**) Characterisation of arsenic sensors in CFS with four different receptor : promoter ratios. The DNA for ArsR expression was 5.3 nM for all the experiments, while the DNA of the promoter circuit was 4.5 nM in **b**, 2.3 nM in **c**, 1.1 nM in **d** and 0.6 nM in **e**. The receptor plasmid was added into cell-free reagent and incubated at 30˚C for 10 min prior to adding the promoter plasmid. All the sensors were induced with various $NaAsO_2$ concentrations as shown in the legends to each graph (ppb for As of $NaAsO_2$). In a 384-well plate, the final volume of cell-free reaction mixture was 50 µl in each well, which was incubated and measured at 37˚C. The plate was sealed with a non-air permeable sealing film. Pos, positive control. a.u., arbitrary units.

5.1.3 Adjusting ArsR-P_{arsR} incubation conditions

As low temperature may improve the protein-promoter binding, I tried different temperatures for the CFS-based arsenic sensor. I first tested the sensor at 30°C instead of 37°C; however, the output expression was slower and no significant improvement of the sensor response was observed (data not shown). I next incubated pre-expressed ArsR with the promoter circuit at 4°C for 10 min prior to the arsenic induction, but no significant improvement was observed (data not shown). In this experiment, ArsR was expressed at 30°C for 10 min in advance, which might not be optimal for ArsR expression; therefore, I next expressed ArsR at 37°C, and then incubated it with the promoter circuit at 4°C. Room temperature was also tested for the receptor-promoter incubation. However, neither of these conditions reduced the sensor background leakage (data not shown).

In conclusion, the arsenic sensor did not work well at low temperature (i.e., 30°C instead of 37°C) which might be due to insufficient transcription and translation. Additional incubation for ArsR-P_{arsR} binding at 4°C or room temperature did not reduce the sensor leakage, neither did it improve the sensor response to arsenic. However, it is not clear whether or not the low temperature improved ArsR binding to P_{arsR}. It is possible that the binding was improved at low temperature, but was quickly reversed during the following 37°C incubation. Nevertheless, whatever the reasons, adjusting the temperature may not be sufficient to solve the leakage issue.

5.1.4 Testing CFS-based arsenic sensors with extra ABS

I first constructed and tested the P_{arsR}-ABS84-*mCherry* circuit (see Section 3.4) in the presence or the absence of the receptor circuit. No significant leakage was observed, however, there was no response to 10 – 100 ppb arsenic induction either (data not shown). Given the promoter strength (i.e., maximum expression level without ArsR repression) should not be decreased by the extra ABS (which was observed in cells, Chapter 3.4), I assumed the commercial cell extract (from L1020, Promega) already contained a sufficient amount of ArsR to strongly repress P_{arsR}-ABS84. Therefore, I increased the amount of P_{arsR}-ABS84-*mCherry* DNA to improve its sensitivity to arsenic induction, and also tested P_{arsR}-ABS67 (with the same concentration) which should be less repressed by ArsR (**Fig. 5.3a**). As expected, P_{arsR}-ABS67 was leakier than P_{arsR}-ABS84 but less leaky than the wild type P_{arsR}. Fortunately, significant responses were observed for 10 ppb arsenic induced P_{arsR}-ABS67 and for 100 ppb arsenic induced P_{arsR}-ABS84. However, P_{arsR}-ABS67 was still very leaky, and the sensitivity of P_{arsR}-ABS84 was very low.

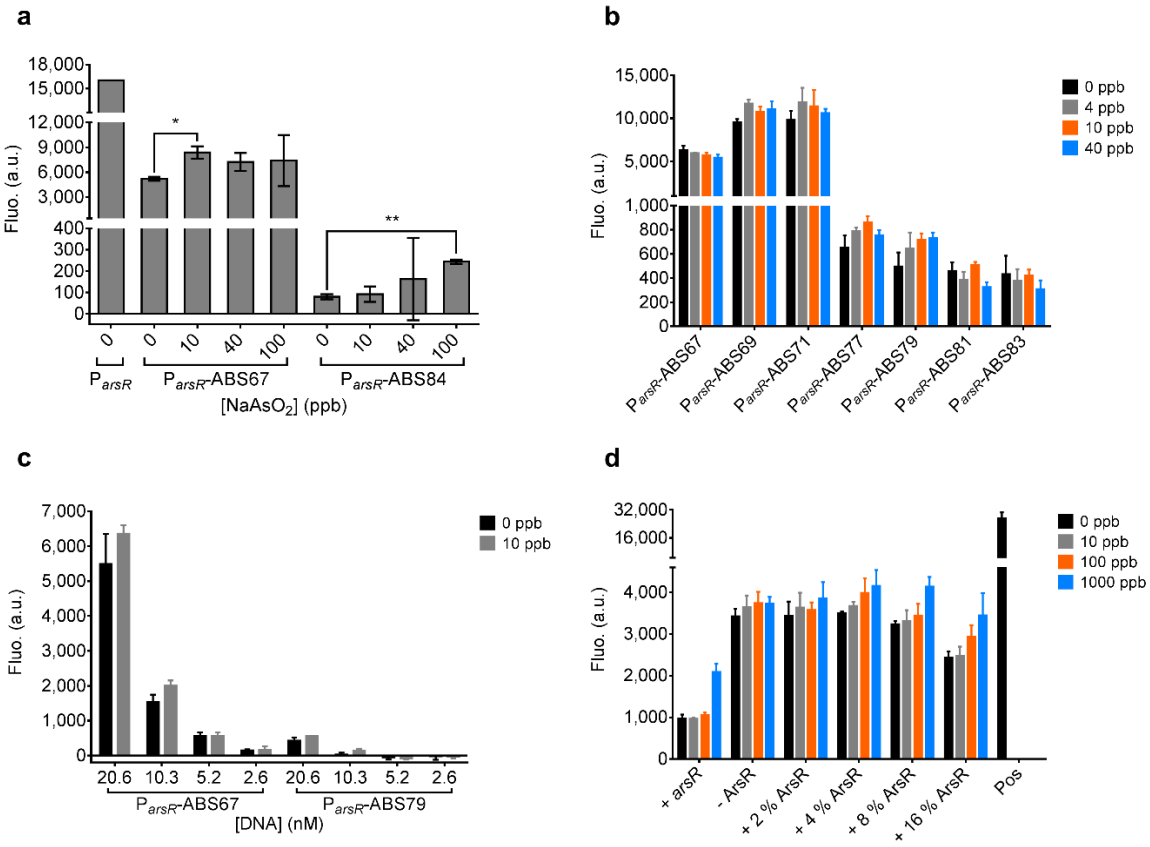


Figure 5.3: Tuning arsenic sensor background leakage in CFS by adding an extra ABS to P_{arsR}.

(**a**) Bar graphs showing the response of arsenic sensors P_{arsR}-ABS67 and P_{arsR}-ABS84 to various concentrations of $NaAsO_2$. For P_{arsR}, no arsenic was added (n = 1). DNA amount for each sensor was 10.8 nM. *, $p \leq 0.05$, **, $p \leq 0.01$ (two-tailed t test). Error bars, s.d. (n = 2, technical replicates). (**b**) Characterisation of seven arsenic sensors with double ABS in CFS. The sensors' DNA were 20.6 nM, and they were induced with 0 ppb, 4 ppb, 10 ppb and 40 ppb As of $NaAsO_2$. Error bars, s.d. (n = 3, technical replicates). (**c**) Reducing the background leakage of P_{arsR}-ABS67 and P_{arsR}-ABS79 by tuning the amount of sensor DNA. The sensors were induced with 0 ppb and 10 ppb $NaAsO_2$. Error bars, s.d. (n = 3, technical replicates). (**d**) Characterisation of P_{arsR}-ABS67 (9.2 nM) in the presence of ArsR expression. +*arsR*, cell-free reaction mixture with 9.2 nM of receptor plasmid. –ArsR, cell-free reaction mixture without receptor plasmid and pre-expressed ArsR. +ArsR, cell-free reaction mixture with pre-expressed ArsR which was pre-expressed in a separate cell-free mixture for 2 h. % represents v/v. Pos, mCherry expression from 40 ng/µl of wild type P_{arsR} plasmid. Error bars, s.d. (n = 3, technical replicates). Time-course characterisation of each sensor is shown in **Appendix Figure 5.1**. The final volume of cell-free reaction mixture was 10 µl for **a**, 5 µl for **b** and **c**, and 4 µl for **d**. Data were collected after 4 h (**a**–**c**) or 8 h (**d**) incubation. a.u., arbitrary units.

To achieve a sensing performance between P_{arsR}-ABS67 and P_{arsR}-ABS84 (i.e., less leaky than P_{arsR}-ABS67 but more sensitive than P_{arsR}-ABS84), I designed and tested multiple arsenic sensors with various distances between the two ABS from 67 bp to 84 bp (**Fig. 5.3b**). Unfortunately, these sensors were still leaky and were not sensitive enough. Some of the sensors were further optimised by adjusting the amount of DNA (**Fig. 5.3c**) and the addition of ArsR (either by adding receptor plasmid or by adding pre-expressed ArsR) (**Fig. 5.3d**), but their sensing performance was still not satisfactory.

According to the results acquired so far, further optimisation could be done to improve the arsenic sensor in CFS. I have shown that reducing the amount of DNA reduced the sensor leakage in CFS (**Fig. 5.3c**), and so did addition of pre-expressed ArsR (**Fig. 5.3d**). Moreover, adding extra ArsR improved the sensor response to arsenic. This might be the best way as it does not require resource from the CFS, and there is no continuous ArsR expression to prevent P_{arsR} from activation. However, if the sensing promoter is too leaky, more pre-expressed ArsR will be required (e.g., for P_{arsR}-ABS67). Although the leakage can be further minimised by reducing the amount of DNA, the optimisation procedure will be too complicated. Therefore, supplying highly pre-expressed ArsR or purified ArsR seems preferable for further optimising the leaky promoters.

It is unclear if P_{arsR}-ABS77–84 was less leaky due to a more efficient roadblocking effect or weaker promoter strength. Theoretically, they should have the same strength as the wild type P_{arsR}, but the current results do not seem to support this (**Fig. 5.3b**). This may be confirmed by inducing the sensors with a high level of arsenic if it is not toxic to the CFS.

5.1.5 Testing CFS-based arsenic sensors with crowding agents

I hypothesised that the weak ArsR-P_{arsR} binding in CFS may due to the insufficient macromolecular crowding effect which may be caused by the lack of cell membranes. Also, it has been shown that crowding agents can either improve or prevent macromolecular interaction and therefore regulate gene expression in CFS; such effects may also depend on the size and the shape of the crowding agents (Ge et al. 2011; Tan et al. 2013; Kuznetsova et al. 2014; Li et al. 2014; Fritz et al. 2015; Shahid et al. 2017). Thus, I tested different inert crowding agents in CFS: PEGs (i.e., polyethylene glycols, mesh-like) with two different molecular weights (i.e., 4 kDa and 8 kDa), and Ficoll 400 (i.e., copolymer of sucrose and epichlorohydrin, spherical, 400 kDa) (**Fig. 5.4**, **Appendix Fig. 5.2**).

Unfortunately, both PEG and Ficoll at low concentrations appeared to accelerate the output expression from P_{arsR} (no matter with or without arsenic induction) rather than improving ArsR-P_{arsR} binding, while both at high concentrations led to low output expression. According

to previous studies (Tan et al. 2013; Shahid et al. 2017), the decreased output expression was likely due to the reduced diffusion of all the macromolecules in CFS rather than improved ArsR-P_{arsR} binding. The lack of obvious effect on ArsR-P_{arsR} binding may be caused by insufficient ArsR present in CFS. Therefore, the experiment should be repeated with extra ArsR in the future.

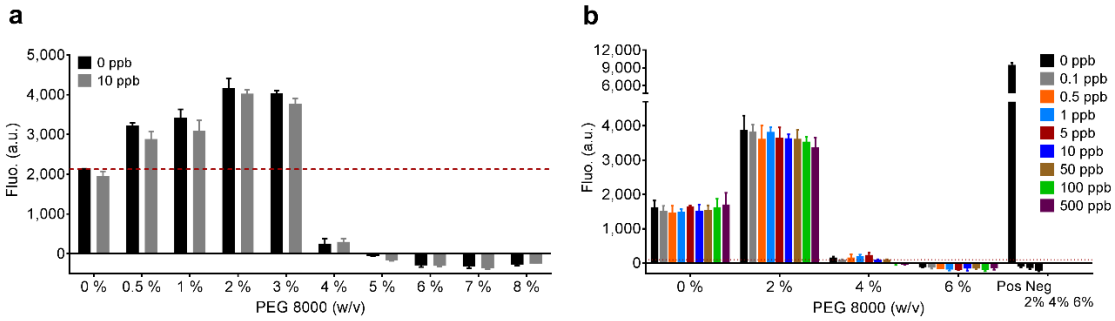


Figure 5.4: Characterisation of an arsenic sensor with crowding agents in CFS.

(**a** and **b**) Characterisation of 9.2 nM P_{arsR}-ABS67 arsenic sensor with PEG 8000 in CFS. The sensor was induced with various concentrations of $NaAsO_2$, which are shown as legends in each graph. The results of other repeated experiments are shown in **Appendix Figure 5.2**. The final volume of cell-free reaction mixture was 5 µl. All the data were collected after 4 h incubation. Pos, mCherry expression from 40 ng/µl of wild type P_{arsR} plasmid. Neg, 40 ng/µl empty plasmid mixed with 2%, 4% or 6% PEG 8000. Error bars, s.d. (n = 3, technical replicates). a.u., arbitrary units.

Further tests were done for PEG 8000 as it initially seemed to improve the response of an arsenic sensor to arsenic (**Appendix Fig. 5.2a,b,d,f,g,k**); however, repeated experiments failed to produce the same results (**Fig. 5.4**, **Appendix Fig. 5.2c,e,h–j,l–p**). The initial positive results may be an artefact due to poor mixing of PEG 8000. The PEG 8000 stock solution was very thick and viscous, and difficult to mix with cell-free reagents. As a result, assays with different levels of arsenic may also have contained different amounts of PEG. The one with 0 ppb arsenic induction level may contain more PEG thus had low output expression (because high PEG improves ArsR binding or reduces the resource diffusion), whereas the one with 10 ppb induction may contain low PEG therefore gave high output expression (because of the sensor leakage, and low PEG improved the transcription/translation).

In summary, macromolecules can either improve or reduce gene expression in CFS. However, PEG 4000/8000 and Ficoll 400 did not improve the sensor response under the tested conditions. Moreover, their effect on ArsR-P_{arsR} binding is still not clear, and existing ArsR in CFS might

not be sufficient. In the future, it will be better to add extra ArsR into the CFS while testing the arsenic sensors with crowding agents.

5.1.6 Engineering arsenic promoters with various σ^{70} binding affinities

In P_{arsR} I used here, the ABS overlaps with a strong –35 region (**Fig. 5.5a**), which might be the reason for low ArsR-P_{arsR} binding affinity in CFS (although it is unclear why this was not a significant issue in cells). A recent study showed that engineering σ^{70} binding sites in a ligand-inducible promoter regulated its background leakage and dynamic range (Chen et al. 2018). Inspired by this study, I designed a small library of arsenic responsive promoters with various –10 and –35 regions. Here are two design strategies (**Fig. 5.5**):

1) In the wild type P_{arsR}, the ABS contains a strong –35 region, but the –10 region is very weak. So, the first strategy is to reduce the strength of the –35 region, and to slightly increase the strength of the –10 region (**Fig. 5.5a**). Based on the library developed by previous studies (Cox et al. 2007; Chen et al. 2018), I selected and designed five –35 regions, and selected five –10 regions. Altogether, 25 combinations could be generated for P_{arsR}.
2) The second strategy is to generate new promoters by merging the ABS with weak –10 and –35 regions (**Fig. 5.5b**). However, the ABS is much longer than a general spacer between –10 and –35, so it needs to overlap with the –10 and –35 regions. The advantage of the overlap is to improve ArsR repression, but it also increases the difficulty of designing the promoters. I identified two partial or full –10 regions in the ABS, and manually added a weak –35 region in front of the ABS to make an entire promoter. In total, 5 promoters were designed.

By applying these two strategies, I aimed to regulate the ArsR-P_{arsR} binding affinity and the promoter leakage, which may lead to a better arsenic response for the CFS-based arsenic sensors. In total, 29 promoters were correctly constructed and tested.

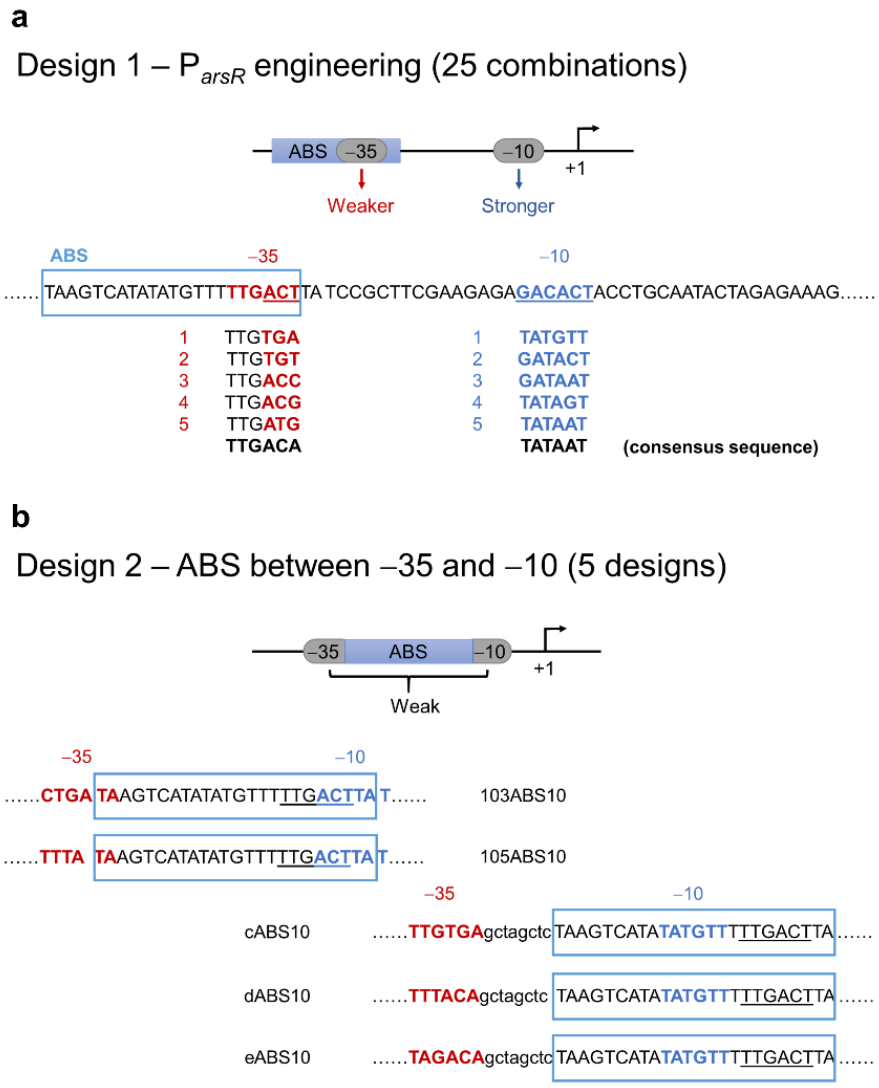


Figure 5.5: Designs of arsenic inducible promoters.

(**a**) Diagram and sequence details of design 1: directly engineering –35 and –10 regions in P_{arsR}. The engineered P_{arsR} will be named after selected –35 and –10 sequences (e.g., $P_{arsR}12$ represents the P_{arsR} with –10 #1 and –35 #2). (**b**) Diagram and sequence details of design 2: engineering arsenic inducible promoters with an ArsR binding site (ABS) overlapping with –35 and –10 regions. The names of engineered promoters are shown next to their sequences.

5.1.6.1 Testing engineered arsenic promoters in cells

I first tested the 29 engineered arsenic promoters in *E. coli* as it is easier and cheaper to do than in CFS. It also can determine the promoter strength, leakage and binding affinity to ArsR (i.e., repression fold of ArsR to the promoter, higher repression fold represents higher binding

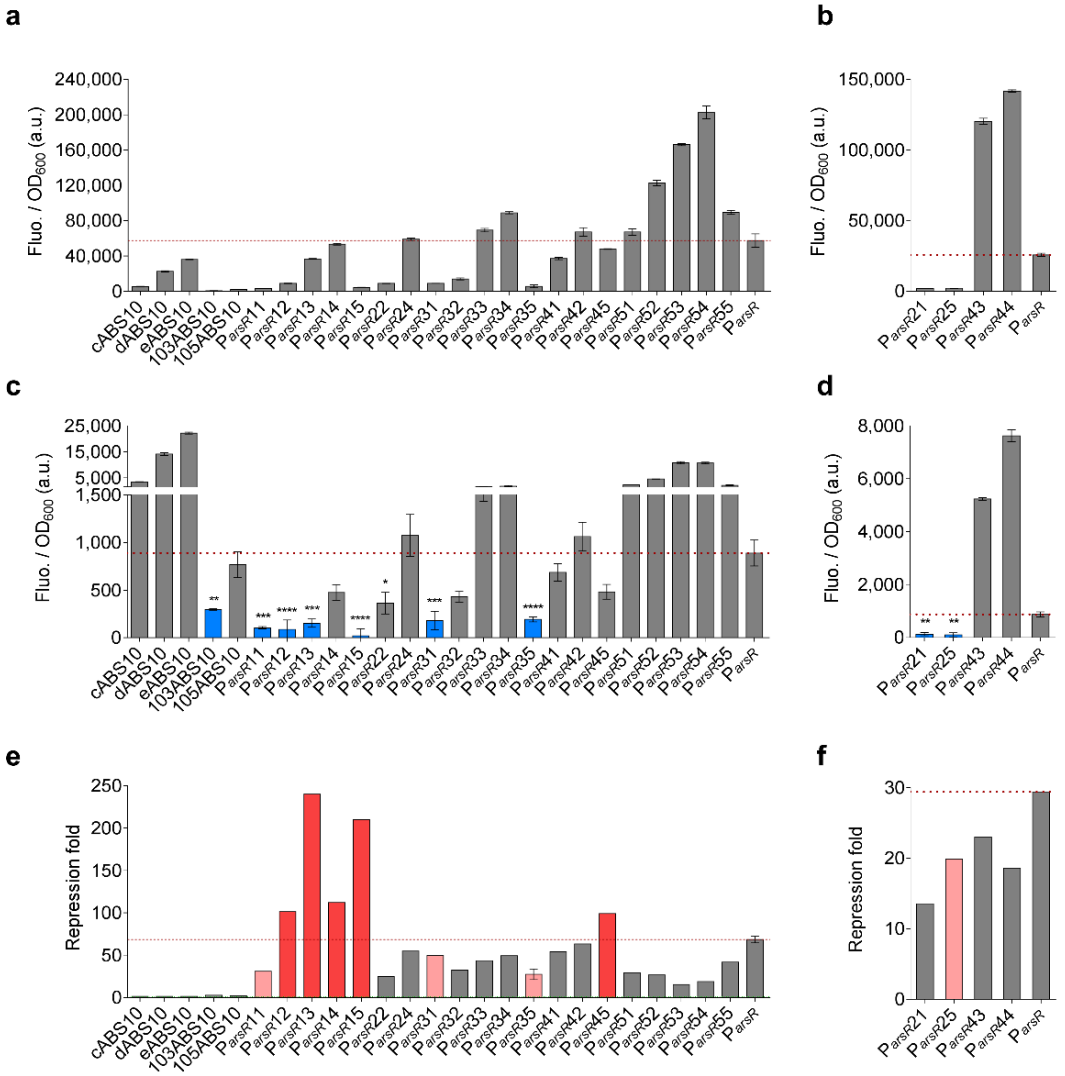


Figure 5.6: Characterisation of engineered arsenic inducible promoters in cells.

(**a** and **b**) *In vivo* characterisation of engineered promoters carried by pSB1A3 (**a**) or pSB3K3 (**b**), showing the promoters' basal expression in the absence of ArsR expression. (**c** and **d**) *In vivo* characterisation of engineered promoters carried by pSB1A3 (**c**) or pSB3K3 (**d**), showing the promoters' basal expression in the presence of ArsR expression. ArsR expression was from the receptor plasmids either on pSB3K3 (**c**) or on pSB4A3 (**d**). Detailed plasmid maps are shown in **Appendix Figure 5.3a** and **c**. Blue bars show the promoters with basal expression significantly lower than the wild type P_{arsR}, with $p \leq 0.01$ (one-way ANOVA test). To show the induction dynamics of the promoters, they were also tested with various $NaAsO_2$ induction, and detailed response is shown in **Appendix Figure 5.3b** and **d**. (**e** and **f**) Repression fold of the engineered promoters repressed by ArsR. It was calculated using the data from **a** – **d**: Repression fold = (Promoter's basal expression without ArsR) / (Promoter's basal expression with ArsR). Promoters highlighted in red and pink are the ones with good repression fold and low background leakage. Red: higher repression than wild type P_{arsR}. Pink: lower repression than wild type P_{arsR}. All the data were collected after 5 h incubation. Error bars, s.d. ($n \geq 2$). a.u., arbitrary units.

affinity). Most of the promoters were ligated to reporter *mCherry* in a high copy number plasmid (i.e., pSB1A3) (**Appendix Fig. 5.3a**), while the others were engineered in a medium copy number plasmid (i.e., pSB3K3) due to cloning issues caused by cell burden (**Appendix Fig. 5.3c**).

To test the promoter strength (and also the basal expression in the absence of ArsR expression), I transformed the promoter plasmids alone into *E. coli*, and measured the *mCherry* expression without arsenic induction (**Fig. 5.6a,b**). To test the promoter leakage in the presence of ArsR expression, I co-transformed the promoter plasmids along with receptor plasmids into *E. coli* (**Appendix Fig. 5.3a,c**), and measured the *mCherry* expression without arsenic induction (**Fig. 5.6c,d**). To estimate the promoter binding affinity to ArsR, I calculated the repression fold using the values from the first two steps: Repression = (Promoter's basal expression without ArsR) / (Promoter's basal expression with ArsR) (**Fig. 5.6e,f**). Promoters with good repression and low leakage were selected for testing in CFS (shown as red and pink bars in **Fig. 5.6e,f**). All the selected promoters responded to arsenic induction *in vivo* (**Appendix Fig. 5.3b,d**).

5.1.6.2 Testing engineered arsenic promoters in CFS

I next tested the selected promoters in CE-CFS (**Fig. 5.7**). Similar to the *in vivo* test, the selected arsenic promoters were less leaky than the wild type P_{arsR} in CFS (**Fig. 5.7b**), and some of them seemed to have better response to arsenic (**Fig. 5.7c**). However, some promoters were still leaky, and the detection limits of all the promoters were above or around 10 ppb arsenic, which was not ideal for applications.

To reduce the background leakage or to improve the sensitivity of the engineered promoters, I proceeded to adjust the ArsR : Promoter ratio and demonstrated this using $P_{arsR}14$ and $P_{arsR}25$. $P_{arsR}14$ is one of the leaky promoters. For this promoter, I adjusted the ArsR : Promoter ratio in two ways: **1)** adding pre-expressed ArsR into the cell-free reaction mixture at different v/v percentages, and **2)** decreasing the amount of the promoter plasmid (**Fig. 5.8a**). The results suggest that the pre-expressed ArsR was not enough to repress $P_{arsR}14$, while reducing the promoter plasmid concentration significantly reduced the background leakage; however, the sensor response to arsenic was also decreased in the second case while the first case seemed to improve the sensitivity. The sensitivity might be difficult to improve in the second case, but the first case might be further optimised by adding more pre-expressed ArsR.

Similarly, I tried to improve the sensitivity of the $P_{arsR}25$-based sensor via reducing the concentration of the receptor plasmid (**Fig. 5.8b**). Unfortunately, the background leakage was also increased; moreover, the sensitivity seemed even to be decreased. These may because ArsR expression was already not sufficient to repress $P_{arsR}25$ before the adjustment; also,

decreasing the receptor plasmid concentration may have increased the resource allocation to the promoter plasmids, which further increased the basal expression and obscured the response to arsenic. Further optimisation might be done by adding extra pre-expressed ArsR to reduce the leakage and to boost the sensitivity.

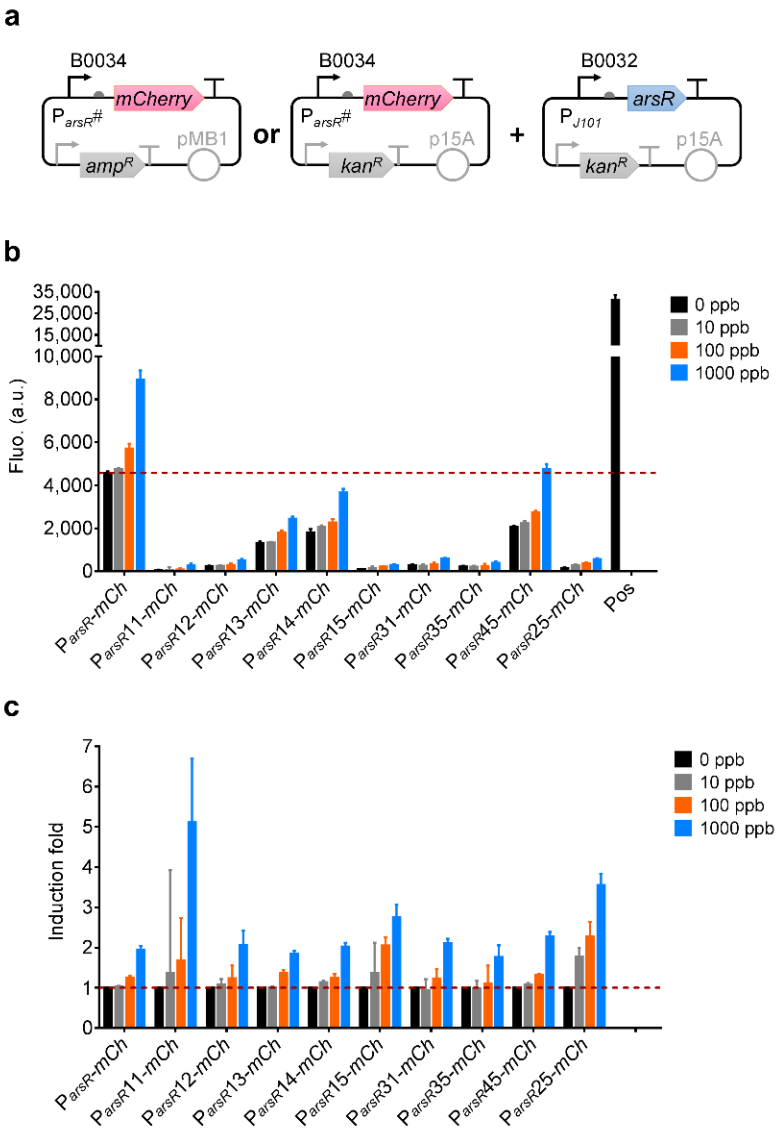


Figure 5.7: Characterisation of engineered arsenic inducible promoters in CFS.

(**a**) Diagrams of genetic circuits tested for **b** and **c**. 8.2 nM of each plasmid DNA was used for all the experiments. (**b**) Characterisation of engineered arsenic inducible promoters in response to 0 ppb, 10 ppb, 100 ppb and 1000 ppb As of $NaAsO_2$ in CFS. Time-course characterisation of each sensor is shown in **Appendix Figure 5.4**. (**c**) Induction fold of the engineered promoters characterised in **b**. The final volume of cell-free reaction mixture was 4 µl in each well. All the data were collected after 8 h incubation. Pos, mCherry expression from 40 ng/µl of wild type P_{arsR} plasmid. Error bars, s.d. (n = 3, technical replicates). a.u., arbitrary units.

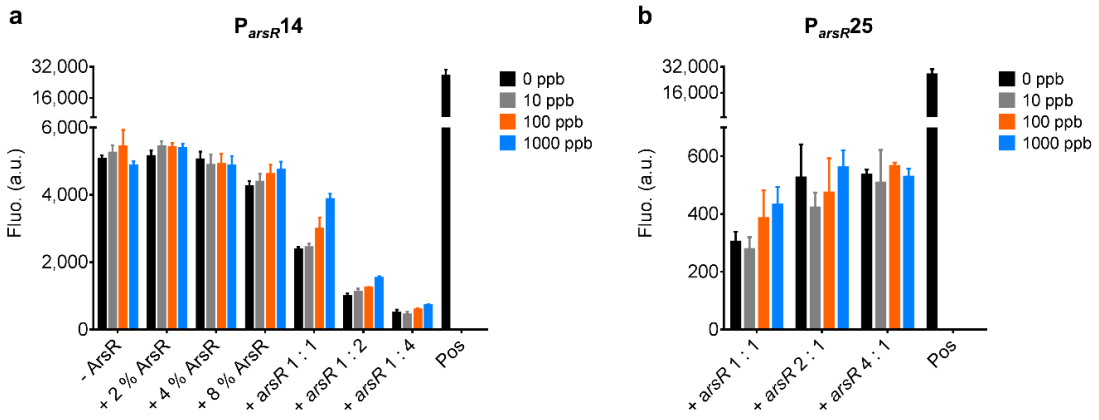


Figure 5.8: Tuning the background leakage and detection limit of engineered arsenic sensors in CFS.

(**a**) Characterisation of P_{arsR}14 in the presence of ArsR expression. –ArsR, cell-free reaction mixture without receptor plasmid and pre-expressed ArsR. +ArsR, cell-free reaction mixture with pre-expressed ArsR which was pre-expressed in a separate cell-free reaction mixture for 2 h. % represents v/v. +*arsR*, cell-free reaction mixture with receptor plasmid, the ratio represents promoter : receptor. P_{arsR}14 was 8.2 nM for adding pre-expressed ArsR. For increasing *arsR* to P_{arsR}14 ratio, the receptor plasmid was always 8.2 nM while P_{arsR}14 DNA concentration was decreasing. (**b**) Characterisation of P_{arsR}25 in the presence of ArsR expression. P_{arsR}25 DNA concentration was always 8.2 nM, while the amount of the receptor plasmid was decreasing. Each sensor was induced with 0 ppb, 10 ppb, 100 ppb and 1000 ppb As of $NaAsO_2$. Time-course characterisation of each sensor under different conditions is shown in **Appendix Figure 5.5** and **5.6**. The final volume of cell-free reaction mixture was 4 µl in each well. All the data were collected after 8 h incubation. Pos, mCherry expression from 40 ng/µl of wild type P_{arsR} plasmid. Error bars, s.d. (n = 3, technical replicates). a.u., arbitrary units.

Although the current optimisation for the engineered promoter-based sensors has not been successful, it is still possible to improve the sensors by changing the output to a more sensitive enzymatic reaction (e.g., LacZ based colorimetric output), especially for the ones which did not show obvious background leakage (e.g., P_{arsR}11/15/25); also, directly wiring the less leaky promoters with amplifiers may boost the sensitivity and the output at low arsenic induction.

I also noticed that none of the promoters from the second design were significantly repressed by ArsR. This suggests the failure of the promoter design, which might be due to ineffective ABS. Based on a previous study (Xu et al. 1996), the adjacent sequences of the current ABS may be involved in ArsR binding. Including these sequences in the promoter design may improve the promoter binding affinity to ArsR.

5.2 CFS-based mercury biosensors

In parallel to optimise CFS-based arsenic biosensors, I tested previous studied mercury sensors in CE-CFS.

5.2.1 Successful initial test of mercury sensors in CFS

The J115-based mercury sensor with GFP/RFP reporter was first tested in CFS, showing better response with RFP than GFP (**Fig. 5.9a,c–f**). Unlike CFS-based arsenic sensors, the background leakage of the mercury sensor was negligible. Moreover, the sensor responded to 2 ppb mercury which met the EPA's guideline.

The J117-based mercury sensor was compared to see whether tuning the receptor density could also affect the sensor sensitivity in CFS (**Fig. 5.9a,g,h**). Contrary to the *in vivo* study (**Fig. 5.9b**), the J117-based mercury sensor was less sensitive with lower output dynamic range than the J115-based mercury sensor in CFS.

J115 and J117 have not been directly compared in CFS; however, previous studies on Anderson promoters in CFS suggest similar relative strengths *in vitro* and *in vivo* (Chappell et al. 2013; UNITN iGEM 2015). Therefore, the MerR expression level from J115 should be higher than that from J117 in CFS.

The sensors' behaviour in CFS are reasonable as MerR is technically an activator; more MerR expression should lead to better P_{merT} activation in response to mercury, and so to the sensor sensitivity and output dynamic range (**Fig. 1.1c,e**). However, the repressing feature of MerR still needs to be considered in the case of high MerR expression, which may increase the required mercury concentration and therefore decrease sensor sensitivity (**Fig. 1.4c**). Thus, there should be a threshold of MerR expression; above the threshold (when MerR exceeds level of its DNA binding site), it can be treated as a repressor otherwise as an activator for the receptor density tuning. Accordingly, the current *in vivo* cases could be considered to be above the threshold while the *in vitro* cases were below the threshold. A mathematical model might be useful to explain this case.

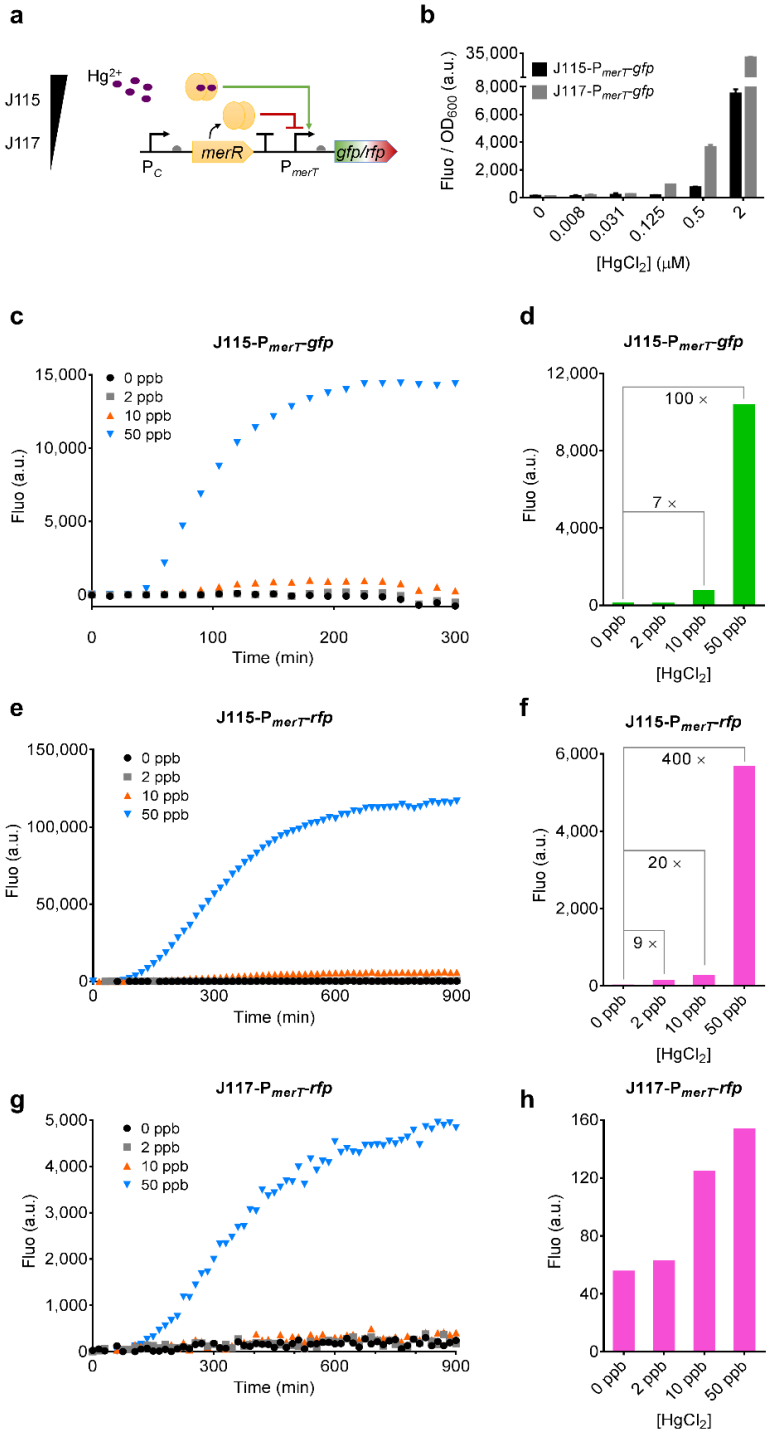


Figure 5.9: Testing mercury sensors with different MerR densities and reporters in CFS.

(**a**) Diagram showing the genetic circuit used in this experiment. (**b**) Characterisation of J115/117-based mercury sensors in cells. Error bars, s.d. (n = 2). (**c** and **d**) Characterisation of J115-based mercury sensor with GFP as reporter in CFS. (**e** and **f**) Characterisation of J115-based mercury sensor with RFP as reporter in CFS. (**g** and **h**) Characterisation of J117-based mercury sensor with RFP as reporter in CFS. 6.7 nM DNA was used for each sensor characterisation. All the sensors were induced with 0 ppb, 2 ppb, 10 ppb and 50 ppb Hg of $HgCl_2$. The final volume of cell-free reaction mixture was 50 μl for each sample. Data for the bar graphs were collected after 2 h incubation. a.u., arbitrary units.

5.2.2 Improving CFS-based mercury sensors by introducing amplifiers

In order to improve the sensitivity and output at low mercury induction, as well as demonstrating the previously studied transcriptional amplifiers (e.g., Amp30^{E}, AmpRinA and AmpE11) in CFS, I rewired these amplifiers with the mercury sensor. All the amplified sensors worked in cells (**Appendix Fig. 5.7a**). The ECF-based amplifiers successfully improved sensor sensitivity and output readout in CFS, and 1 ppb detection limit was achieved (**Fig. 5.10**). Moreover, by tuning the RBS for ECF11 expression, I reduced the sensor leakage and improved the sensor output dynamic range (**Fig. 5.10b,d–g**).

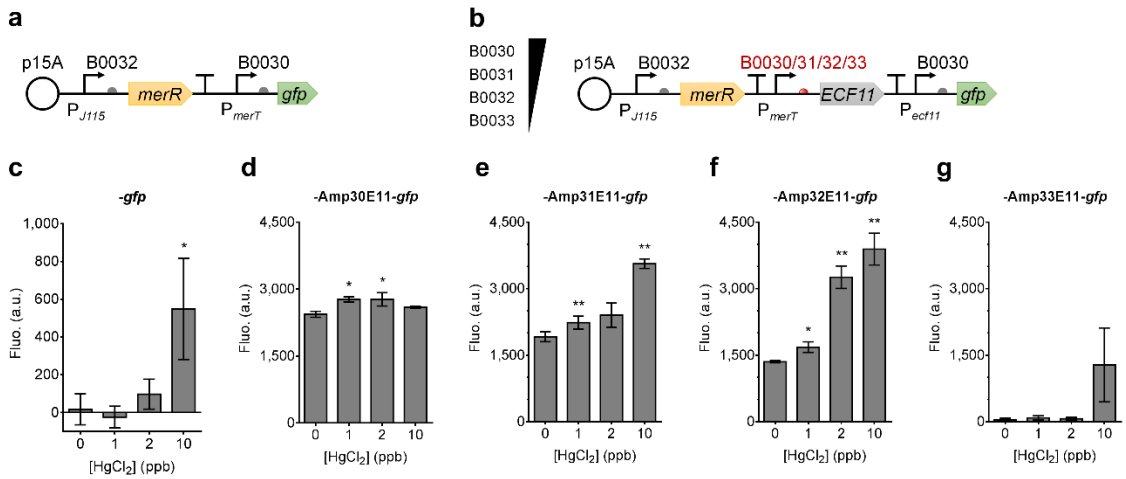


Figure 5.10: Improving sensor sensitivity and output dynamics by employing amplifiers in CFS.

(**a**) Diagram of a mercury sensor characterised in **c**. (**b**) Diagram of a mercury sensor with ECF-based amplifiers characterised in **d**–**g**. The RBS for ECF expression was varied from strong B0030 to weak B0033 (**Appendix Table 2.1**). (**c**–**g**) Characterisation of the mercury sensors without or with ECF-based amplifiers. -Amp30E11 represents the amplifier with RBS30 regulated ECF11. 12.5 nM DNA was used for each sensor characterisation. All the sensors were induced with 0 ppb, 1 ppb, 2 ppb and 10 ppb Hg of $HgCl_2$. *, $p \leq 0.05$, **, $p \leq 0.01$ (two-tailed t test). The final volume of tested cell-free reaction mixture was 5 µl for each sample. Data were collected after 2 h incubation. Time-course characterisations of sensors in **c** and **f** are shown in **Appendix Figure 5.7**. Error bars, s.d. (n = 3, technical replicates). a.u., arbitrary units.

However, the sensors with all the other amplifiers (e.g., Amp30^{E} and AmpRinA) did not respond to mercury in CFS (data not shown). This again implied the intrinsic differences between the cells and CFS. The failure of HrpRS and RinA-based systems could be caused by many reasons, such as insufficient protein folding/maturation of HrpRS and RinA, as well as insufficient transcription from P_{hrpL} and P_{rinA} (maybe due to inadequate components for

supporting these systems in CFS); excess resource allocation to HrpRS expression may also cause insufficient reporter expression.

As the Amp32E11 gave the best result in terms of detection limit and output dynamic range, it was further optimised to reduce its leakage by lowering the sensor DNA concentration in CFS (**Fig. 5.11**). From 12.5 nM to 6.2 nM, the sensor leakage was significantly reduced while maintaining similar induction by mercury. With 3.1 nM DNA, the leakage was negligible, but the detection limit was significantly increased. The sensor leakage might be further optimised by tuning the DNA between 6.2 nM and 3.1 nM or by introducing a protein degradation system.

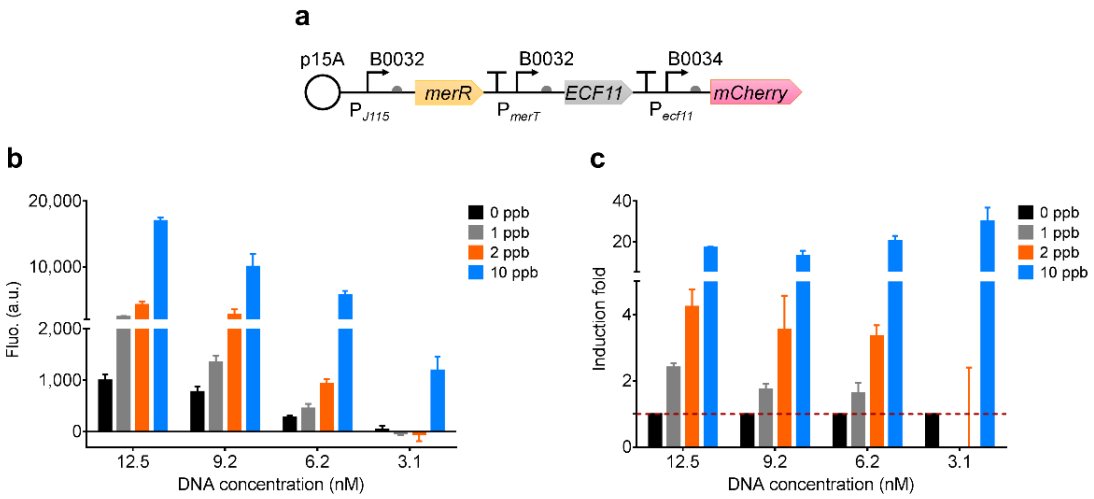


Figure 5.11: Reducing mercury sensor background leakage by tuning DNA concentration in CFS.

(**a**) Diagram showing the mercury sensor used in this experiment. (**b**) Characterisation of the mercury sensor with different DNA concentration in response to 0 ppb, 1 ppb, 2 ppb and 10 ppb Hg of $HgCl_2$. The final volume of cell-free reaction mixture was 5 µl for each sample. Data were collected after 4 h incubation. Time-course characterisation of the sensor is shown in **Appendix Figure 5.8**. (**c**) Induction of the sensor characterised in **b**. Error bars, s.d. (n = 3, technical replicates). a.u., arbitrary units.

I also tested other ECF-based amplifiers trying to identify the ones with less background leakage but high output dynamic range in CFS. ECF16/17/20/22 with different RBS were selected based on my colleague Filipe Pinto's *in vivo* characterisation (**Appendix Fig. 5.9a**, ECF16/17/20/22-based systems were less leaky but with higher output dynamic range than Amp32E11 when characterised using a P_{BAD} promoter). I ligated those amplifiers with the J115-based mercury sensing module, and first tested them *in vivo* (**Appendix Fig. 5.9b,c**). However, the sensing module and the amplifiers were too toxic to the cells upon mercury induction, so their responding performance were not directly comparable (**Appendix Fig. 5.9b,c**). Hence, I characterised these amplifiers in CFS (**Fig. 5.12**).

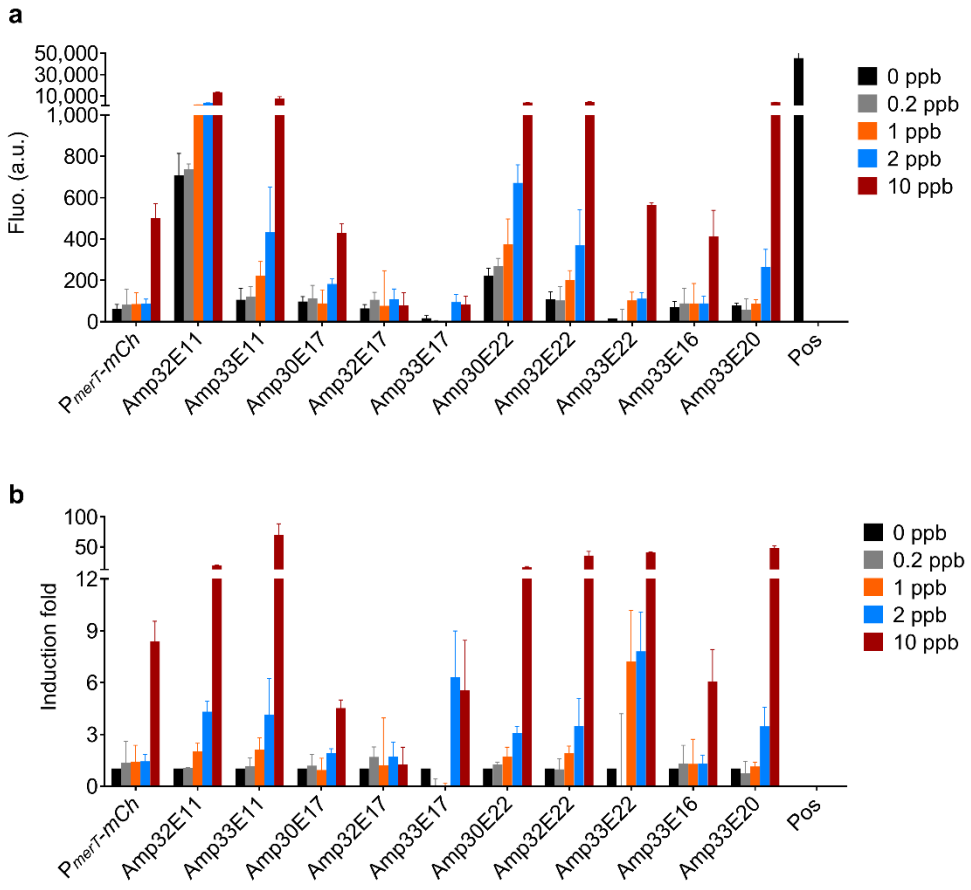


Figure 5.12: Characterisation of ECF-based amplifiers in CFS.

(**a**) Characterisation of a mercury sensor with different ECF-based amplifiers in response to 0 ppb, 0.2 ppb, 1 ppb, 2 ppb and 10 ppb Hg of $HgCl_2$. 6.2 nM DNA was used for each sensor. The final volume of cell-free reaction mixture was 4 µl for each sample. Data were collected after 12 h incubation. (**b**) Induction of the sensors characterised in **a**. Pos, mCherry expression from 40 ng/µl of wild type P_{arsR} plasmid. Time-course characterisation of each sensor is shown in **Appendix Figure 5.10**. Error bars, s.d. (n = 3, technical replicates). a.u., arbitrary units.

All the tested amplifiers had less background leakage than the Amp32E11 but were less sensitive (**Fig. 5.12a**) in CFS. Nevertheless, some of the amplifiers had good output dynamic ranges and they may be cascaded with other amplifiers to achieve better output dynamic range (e.g., Amp30/32E22, Amp33E20 and Amp33E11) (**Fig. 5.12b**). In addition, as some had low basal expression, they might be directly tested with enzyme-based colorimetric output to improve the sensitivity and the output.

Some differences were observed while comparing the data from *in vivo* and *in vitro* experiments. For instance, Amp33E16 and Amp33E20 were supposed to have higher output expression levels than Amp32E11 (**Appendix Fig. 5.9a**); however, the opposite was true in

CFS. It is unclear whether the differences were caused by the different sensing platform, as the full input range of mercury induction might not have been tested.

5.2.3 Testing mercury sensor and macromolecules in CFS

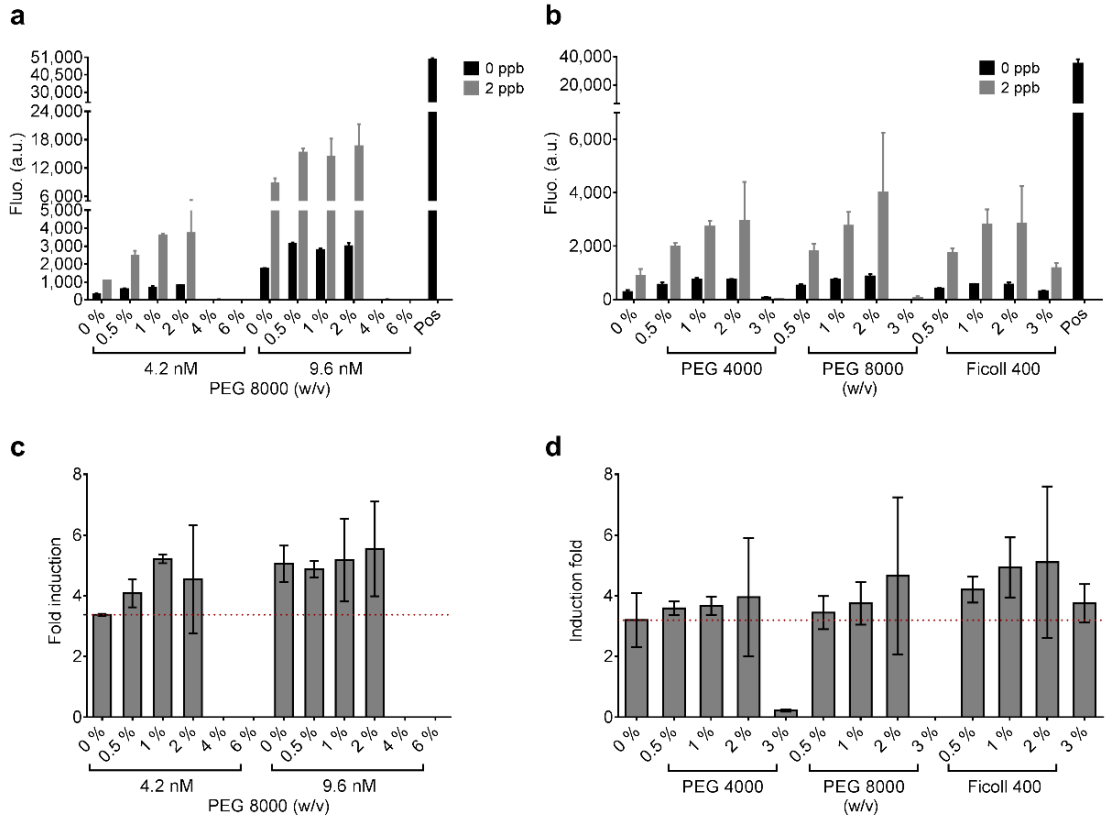


Figure 5.13: Characterisation of a mercury sensor with crowding agents in CFS.

(**a**) Characterisation of a mercury sensor supplied with various percentages of PEG 8000 in CFS. 4.2 nM and 9.6 nM DNA were used. (**b**) Characterisation of a mercury sensor supplied with various concentrations of PEG 4000, PEG 8000 and Ficoll 400 in CFS. 4.2 nM DNA of the mercury sensor (J115-*merR*-P_{merT}-Amp32E11-*mCherry*) was used. The sensor was induced with 0 ppb or 2 ppb Hg of $HgCl_2$. The final volume of cell-free reaction mixture was 5 µl for each sample. Data were collected after 8 h incubation. (**c** and **d**) Induction fold of the sensor characterised in **a** and **b**. Pos, mCherry expression from 40 ng/µl of wild type P_{arsR} plasmid. Error bars, s.d. (n = 3, technical replicates). a.u., arbitrary units.

For the same reason as for the arsenic sensors, I also tested mercury sensors with extra macromolecules in CFS. The mercury sensor with Amp32E11 and mCherry output (**Fig. 5.11a**) was first tested with PEG 8000 in CFS (**Fig. 5.13a,c**). The reason to not use GFP as reporter was because the PEG 8000 interfered with the background fluorescence in CFS (data not

shown). Similar to previous results, low PEG 8000 concentration increased the sensor output expression while it was decreased by the addition of high amount of PEG 8000. Although a minute improvement of induction (i.e., higher induction fold) was observed for the condition with 1% PEG 8000 and 4.2 nM DNA, this could not be repeated in later experiments (**Fig. 5.13b,d**). PEG 4000 and Ficoll 400 were also tested using the same sensor, but still no significant improvement of sensor response was achieved (**Fig. 5.13b,d**).

Overall, the induction of the demonstrated sensor was generally stable regardless of the tested macromolecules, but was dramatically reduced when the PEG was ≥ 3%.

5.2.4 Engineering mercury sensors for paper-based CFS

Before developing sensors in a paper-based platform, I built a homemade humidifying chamber for incubating the paper to avoid evaporation, and verified that the paper-based CFS was functional in it (**Fig. 5.14b,c**). As I aimed to develop sensors with colorimetric output on paper, I rewired the J115-based mercury sensor with LacZ as a reporter (**Fig. 5.14a**). I first tested the background of LacZ in the commercial CE-CFS (L1020, Promega) by incubating it with one of LacZ's substrates, X-gal, overnight (**Fig. 5.14d**). The results indicate that there is LacZ existing in the commercial CE-CFS; however, the background colour change can be delayed and reduced by lowering the substrate concentration.

Next, I tested the mercury sensor with various mercury induction levels in a paper-based CFS supplied with X-gal (**Fig. 5.14e**). Although the sensor responded to 2 ppb mercury, the colour change was very slow and there was strong background colour change in non-induced samples. I also found that the solvent for X-gal, Dimethylformamide (DMF), might be toxic to the CFS (e.g., the sensor response to 10 ppb with 0.1% X-gal was slower than the one with 0.05% X-gal), and this might be the reason for the slow reaction.

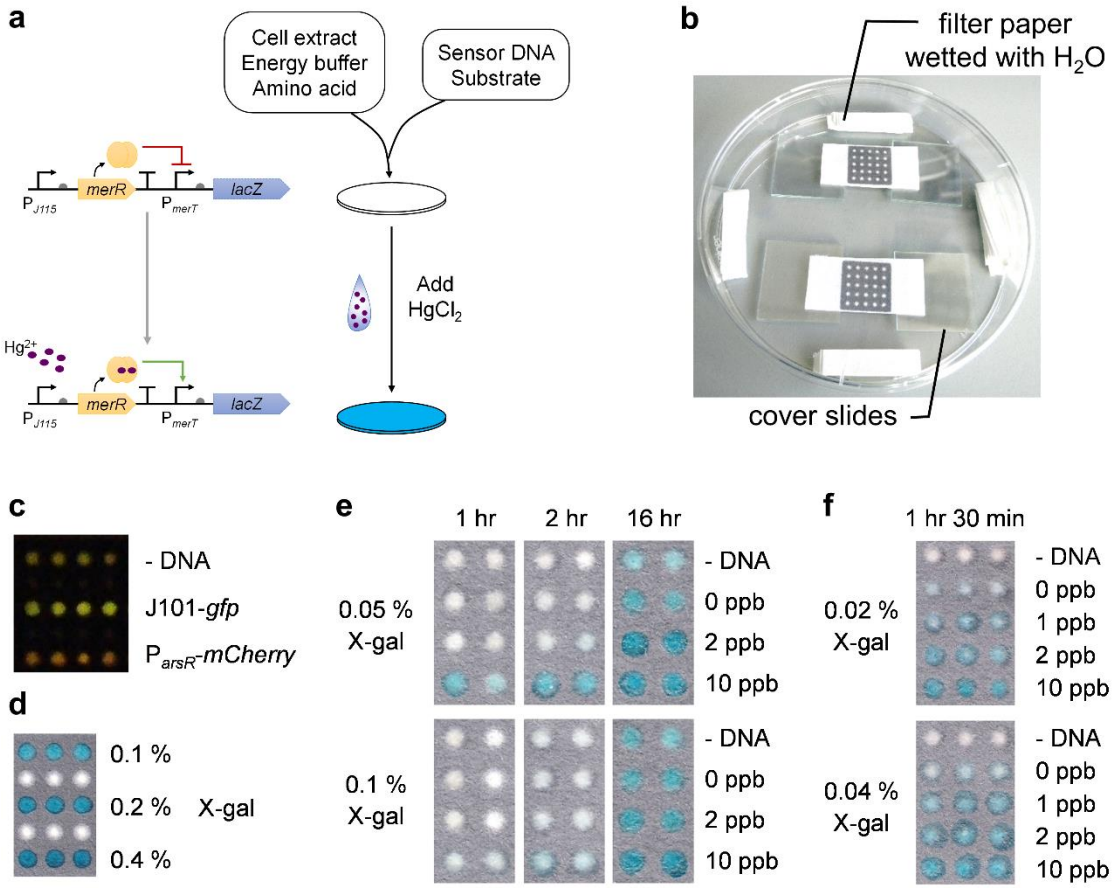


Figure 5.14: Testing a mercury sensor in a paper-based CFS.

(**a**) Diagram showing a mercury sensor with colorimetric output and its testing procedure on paper. (**b**) Image of a homemade humidifying chamber for incubating the paper. The petri dish is sealed with parafilm during incubation. (**c**) Testing paper-based CFS using fluorescent reporters. 40 ng/µl DNA was used for each genetic circuit, and the paper was incubated overnight. - DNA represents the samples without DNA. (**d**) Test of LacZ background in CFS using X-gal as substrate. The paper was incubated overnight. % represents w/v. (**e**) Characterisation of the mercury sensor in the paper-based CFS. 9.7 nM DNA was used for the sensor, which was induced with 0 ppb, 2 ppb and 10 ppb $HgCl_2$. The X-gal was diluted from a 2% stock solution. (**f**) Characterisation of the mercury sensor in the paper-based CFS. 9.7 nM DNA was used for the sensor, and the sensor was induced with 0 ppb, 1 ppb, 2 ppb and 10 ppb $HgCl_2$. The X-gal was diluted from a 5% stock solution. 2 µl of cell-free mixture was spotted on each circle of the paper. Cell phone images were taken following a time course for **f**, and are shown in **Figure 5.15**.

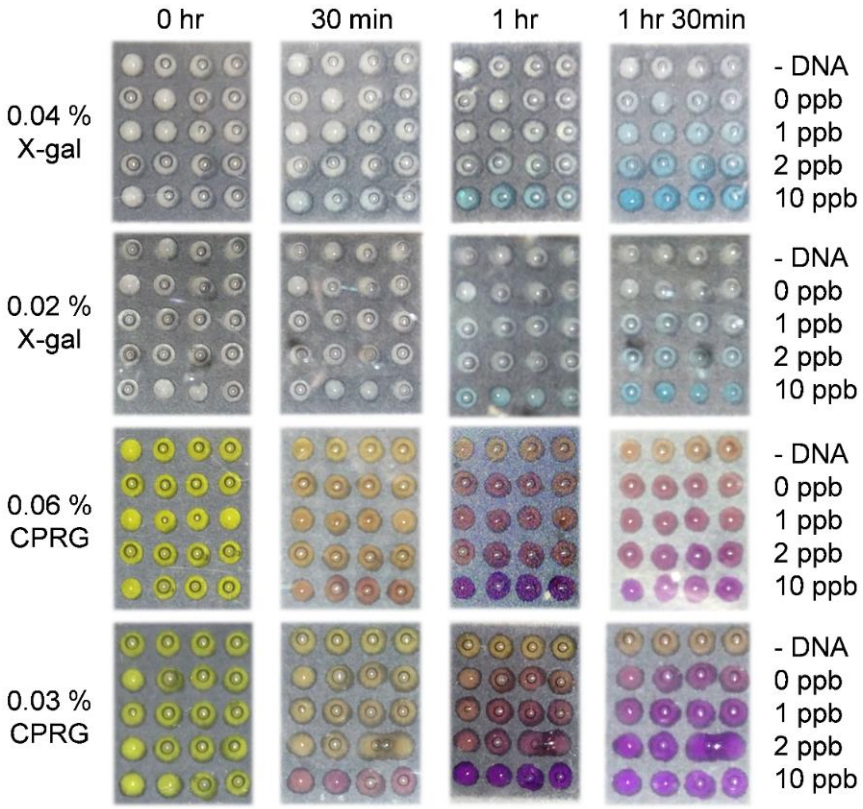


Figure 5.15: Comparing X-gal and CPGR using paper-based CFS.

Cell phone images comparing the two substrates, X-gal and CPRG, of LacZ, as well as recording the experiments shown in **Figure 5.14f**. 0.04% X-gal and 0.06% CPRG are equal to 1 mM.

To reduce the toxicity caused by DMF, I made an X-gal stock solution with higher concentration so that less DMF would be brought into the CFS. With similar X-gal final concentration, less DMF indeed improved the activity of the CFS and accelerated the sensor response (**Fig. 5.14f**). More interestingly, the sensor responded to 1 ppb mercury which is lower than the EPA's guideline. Phone images were also taken over time (**Fig. 5.15**), showing that the sensor response with 0.04% X-gal was faster and more distinct than with 0.02% X-gal. The same molarity of a more sensitive LacZ substrate, CPRG (i.e., chlorophenol red-β-D-galactopyranoside), was also tested, showing its faster response but higher background colour change than X-gal (**Fig. 5.15**).

To reduce the basal expression of the mercury sensor and the background colour change in the paper-based CFS, I tested the system with reduced sensor DNA concentration for the test with X-gal and decreased both DNA and substrate concentrations for the test with CPRG. Unfortunately, the reduced DNA concentration was too low for the sensor to be functional

(e.g., its response to 2 ppb mercury was very slow and not significant) (**Fig. 5.16**). Although a slight sensitivity improvement was observed for 0.015% CPRG with 9.7 nM DNA at 80 min, the background colour change was still significant (**Fig. 5.16b**).

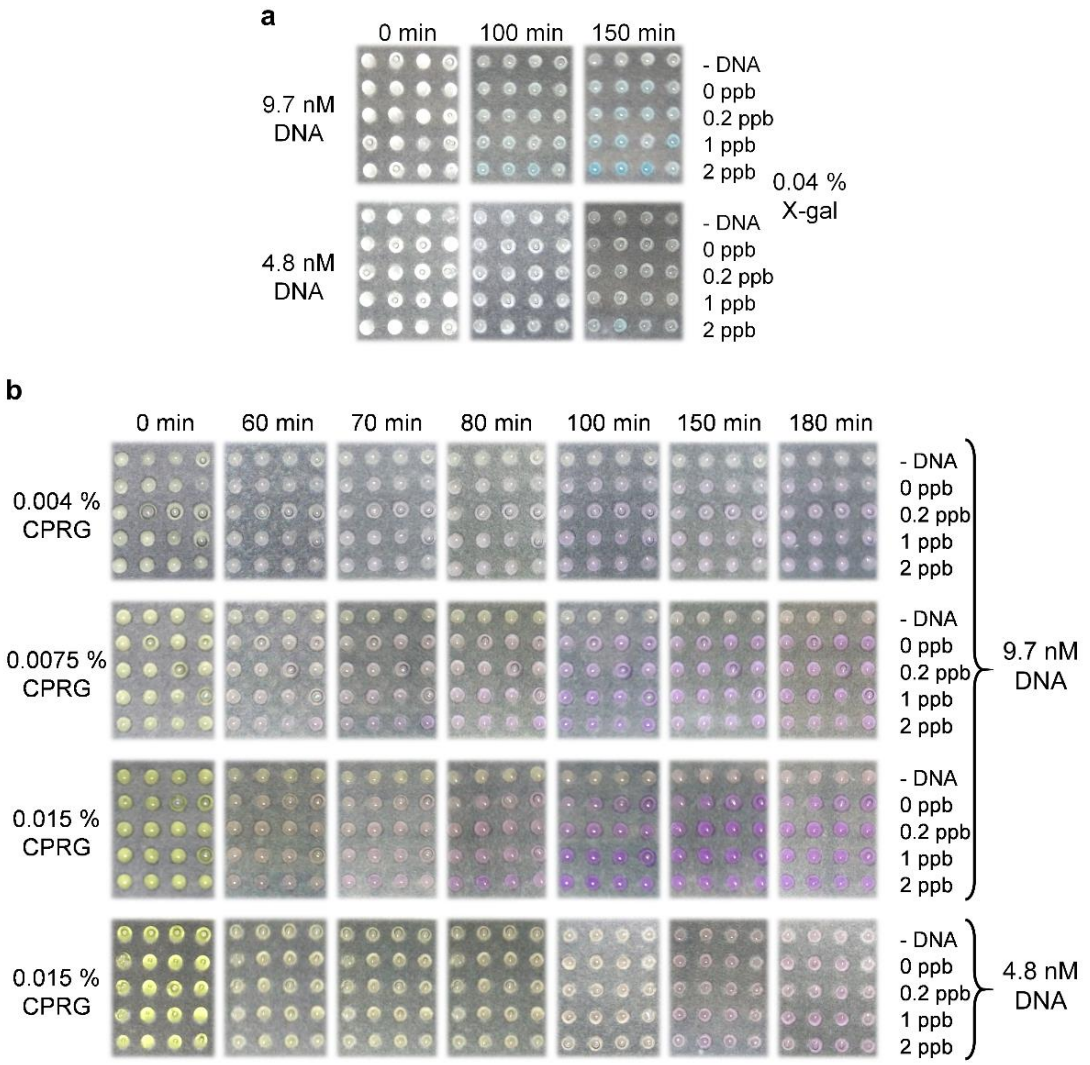


Figure 5.16: Regulating mercury sensor background leakage in a paper-based CFS by tuning the concentration of DNA and substrates.

(**a**) Characterisation of a mercury sensor with reduced DNA on paper. (**b**) Characterisation of a mercury sensor with reduced CPRG and DNA on paper. The sensor was induced with 0 ppb, 0.2 ppb, 1 ppb and 2 ppb Hg of $HgCl_2$. - DNA represents the samples without DNA. 1 µl of cell-free mixture was spotted on each circle of the paper.

Overall, simply reducing the CPRG concentration was not sufficient to reduce the background colour change, which may be due to its high sensitivity to low amounts of LacZ expression;

also, 4.8 nM DNA was not enough to trigger a response to mercury. Thus, it would be better to reduce the background colour change via a more direct way, such as eliminating the background LacZ levels by using LacZ deficient cell extract and introducing a protein degradation system into the sensor.

Another issue I noticed here was that the wax barrier in the chromatography paper (Whatman, 3001-861, from Pardee et al. 2014) might not be sufficient to block diffusion between adjacent samples. Colour diffusion was observed after overnight incubation for some of the CPRG tests, and it was unclear whether it happened in other tests (because some colour diffusion might exist but not be obvious to the naked eye). If it was the case for all the other tests, it might be the reason that the sensor response to 1 ppb or 2 ppb mercury in **Figure 5.16a** was not as obvious as in **Figure 5.14f** and **Figure 5.15**: the strong colour change in 10 ppb induced samples may have diffused to the 2 ppb and 1 ppb induced samples, which made their colour changes more obvious.

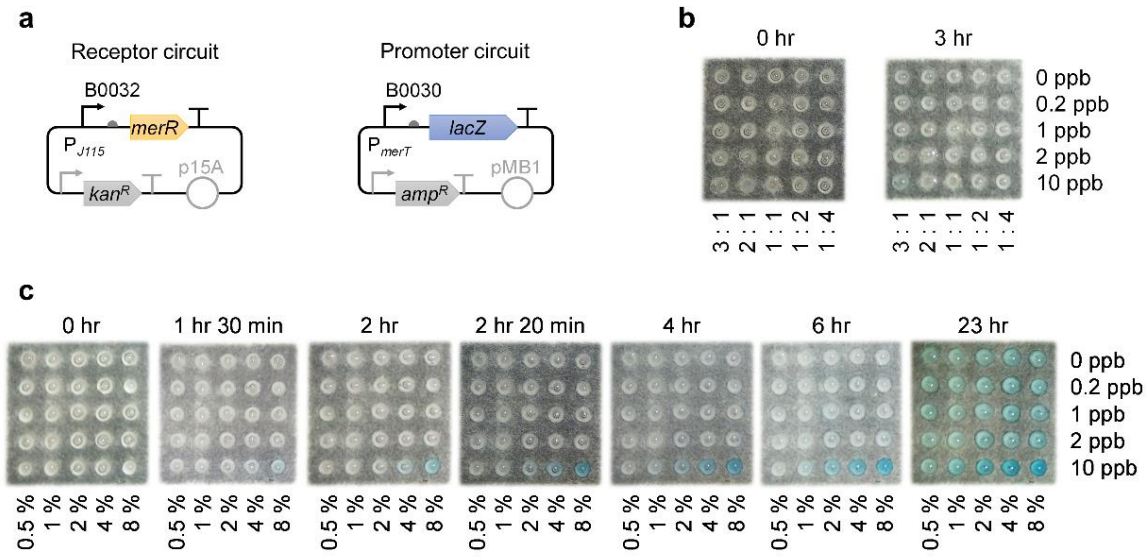


Figure 5.17: Regulating mercury sensor detection limit by tuning MerR : PmerT ratio in a paper-based CFS.

(**a**) Diagrams showing genetic circuits used in **b** and **c**. (**b**) Tuning sensor detection limit by regulating the MerR : P_{merT} DNA ratios. The promoter circuit was fixed at 4.1 nM. (**c**) Tuning the sensor detection limit by regulating the amount of pre-expressed MerR. The promoter circuit was fixed at 9.7 nM. MerR was pre-expressed overnight (for 18 h) from a cell-free reaction containing 40 ng/μl receptor circuit. % represents v/v. The sensor was induced with 0 ppb, 0.2 ppb, 1 ppb, 2 ppb and 10 ppb Hg of $HgCl_2$. 1 μl of cell-free mixture was spotted on each circle of the paper.

Additionally, I tried to generate paper-based mercury sensors with different detection limits. I regulated the detection limit by tuning the MerR : P_{merT} ratio in two ways: **1)** regulating the DNA ratios of the receptor and the promoter circuits (**Fig. 5.17a,b**) and **2)** regulating the amount of pre-expressed MerR in CFS (**Fig. 5.17c**). Unfortunately, this has not been successful so far. Only the sensors with highest MerR and highest mercury induction showed a colour change on paper in short time (90 min), suggesting that the rest of the sensors did not have enough MerR to sufficiently respond to mercury and activate P_{merT}. Accordingly, upregulating the MerR level may lead to better detection limits.

5.3 Discussion

In recent years, CFS have become more and more popular as a biosensing platform (Pellinen et al. 2004; Pardee et al. 2014; Pardee et al. 2016a; Didovyk et al. 2017; Duyen et al. 2017; Wen et al. 2017). This is proposed as a solution to circumvent the biosafety and regulation issues associated with whole-cell biosensors; also, CFS is easy and cheap to produce, and their stability and portability can be extended by freeze-drying on paper (Pardee et al. 2014; Pardee et al. 2016b).

One aim of my PhD project is to develop paper-based arsenic and mercury sensors using CFS, and create 'traffic light' sensor arrays for both systems on paper. I believe these sensing platforms will greatly support the large demand for cheap, easy and safe detection methods for arsenic and mercury.

Although CFS have been demonstrated successfully to support biosensors for sensing heavy metals, pathogens, antibiotics and viral RNA, problems still occurred when transferring the biosensors from cells to CFS in this study: **1)** the arsenic sensors were very leaky, **2)** HrpRS/RinA-based amplifiers did not work and **3)** MerR densities needed to be tuned in the opposite way (to that used in cells) to improve sensor sensitivity. The major cause of those issues is probably the intrinsic differences between the CFS and intracellular environment, such as macromolecular crowding conditions, components for TX-TL and resource finiteness. Additionally, as a reporter in CFS, especially for weak promoters, a red fluorescent protein is better than a green fluorescent protein due to the background fluorescence in CFS.

Because of the issues of the functionality transfer from cells to CFS, it was necessary to re-optimise the sensors in CFS, especially for the sensor background leakage and sensitivity. A couple of methods have been demonstrated to be successful in CFS though still requiring further tuning (major cell-free sensors and their optimised performance are listed in Appendix **Table 5.1**): **1)** regulating the sensor sensitivity by tuning receptor densities, **2)** regulating sensor leakage and output dynamics by tuning receptor : promoter ratios, **3)** improving sensor

sensitivity and output dynamics/readout by introducing a transcriptional amplifier, and **4)** tuning the sensor leakage and output dynamic range by promoter engineering. Receptor densities and receptor : promoter ratios can be tuned either by varying the amount of added pre-expressed receptors or by varying the amount of the receptor/promoter plasmids. However, directly tuning the pre-expressed receptors seems to be a better way than tuning the receptor plasmids. This is because the absence of continuous expression of receptors not only may improve the repressor-based sensor sensitivity, but also leads to a maximum resource allocation to the promoter plasmids and therefore further improves the sensor response.

Addition of inert macromolecules can either improve or reduce sensor output expression, which can be an additional way to regulate sensor leakage; however, they do not seem to improve the sensor's response. Moreover, attention needs to be paid when mixing highly concentrated macromolecules into CFS; poorly mixed samples will dramatically affect the sensor behaviour.

Particularly for the arsenic sensors, the leakage was problematic in CFS, and several tuning methods were tested to address this issue while improving the sensitivity. In addition to the DNA/ArsR regulation, promoter engineering seemed to be a promising way to reduce the basal expression of the arsenic promoter. Fine tuning of the amount of DNA and pre-expressed ArsR, as well as introducing amplifiers and enzyme reporters, might be solutions to improve the sensor sensitivity and output dynamic range. In addition, as insufficient macromolecular crowding might be the cause of the high background leakage of P_{arsR}, the sensor might be improved by mimicking a similar crowding environment as in living cells, such as developing the CFS-based sensors in artificial cells (Tan et al. 2013; Garamella et al. 2016; Adamala et al. 2017; Rampioni et al. 2018) or Simcells (Rampley et al. 2017). Otherwise, testing other arsenic inducible systems and directed evolution, e.g, phage-assisted continuous evolution (PACE) (Esvelt et al. 2011) and compartmentalized partnered replication (CPR) (Abil et al. 2017), might ultimately provide better solutions.

Additionally, I initiated a study on mercury sensors in the paper-based CFS, using LacZ as reporter and X-gal/CPRG as its substrate. So far, the sensor can sense lower than 2 ppb mercury on paper. A couple of issues have been observed: **1)** X-gal's solvent, DMF, was toxic to CFS, but the toxicity can be relieved by using highly concentrated stock solutions; though has not been tested, an alternative solvent, DMSO (dimethyl sulfoxide), may be less toxic. **2)** CPRG was very sensitive to LacZ which accelerated the sensor response but also increased the background colour change, and current attempt to reduce such leakage have not been successful. **3)** There might be sample crosstalk on the paper (due to insufficient wax barrier)

and **4)** the sample loading area was too small even for 1 μl cell-free mix. Also, for tuning the detection limit of the mercury sensor in CFS, current pre-expressed MerR was not sufficient.

To reduce the background colour change, the most immediate way would be using CFS lacking LacZ and performing optimisation (e.g., DNA tuning and protein degradation) based on this system. Other reporters that are not present in *E. coli* cell extract can also be tested and compared.

To achieve various detection limits, purified MerR can be added to the CFS, where the MerR quantity can be controlled easily. In addition, an easier and cheaper way might be using cell extract made from cells expressing different levels of MerR; such extracts can be used for making the CFS or can be added to the CFS as a supplement. Additionally, amplifiers may also contribute to various detection limits in the paper-based CFS.

To eliminate the cross talk on paper, two approaches are proposed here. One approach is to punch out the paper discs and place them separately for testing (e.g., in a 384-well plate or on a chip) (Pardee et al. 2014). Although increasing work load, this way can ensure that there is no crosstalk between adjacent samples. Another approach is to optimise the paper layout and wax melting procedure. This can also solve the problem of small loading area. However, this will require a wax printer, a 120˚C hot plate and may require some test-optimisation cycles (Carrilho et al. 2009). Nevertheless, for point-of-care use, the second method will be necessary as it will be cheap to produce and easy-to-use for large-scale sample screening.

Chapter 6. Summary and Future Directions

Whole-cell biosensors have drawn increasing attention as alternatives to traditional chemical and physical detection methods. Numerous whole-cell biosensors have been developed to sense and report a variety of targets. Although very few have been commercialised due to their poor sensing performance and biosafety concerns, with the help of synthetic biology, such biosensors can be further improved to meet their application requirements.

Here, I proposed and demonstrated an innovative and comprehensive methodology, which can be easily and rapidly applied for improving many whole cell-based biosensors in terms of detection limit, output readout and basal expression levels (Chapter 3). Moreover, I developed easy-to-use sensing platforms which can facilitate the use of the cellular biosensors in the field and mitigate the biosafety issues (Chapter 4). In addition, I adapted the cellular biosensors to a cell-free system (CFS) which can further address the biosafety and regulation concerns (Chapter 5). A number of optimisation methods were also proposed and tested to improve the sensors in CFS.

The sensor optimisation methodology comprises several synthetic biology strategies: **1)** tuning receptor densities to improve sensor sensitivity, **2)** introducing a single transcriptional amplifier or multi-layered amplifier cascade to boost sensor output readout, **3)** engineering promoter architecture and **4)** regulating post-translational reporter protein degradation to reduce sensor background expression. This methodology along with other relevant strategies are summarised in **Fig. 6.1**, and different optimisation methods should be selected based on the knowledge of sensing mechanisms and sensors' performance during optimisation. The first two strategies have been successfully applied to both arsenic and mercury sensors, and their detection limits and outputs were improved up to 5,000-fold and 750-fold respectively. As no significant leakiness was observed in mercury sensors, the last two strategies were only tested for arsenic sensors and they successfully reduced the sensors' background leakage. Despite the success of each strategy, several caveats were raised during this research and are summarised below.

For the first strategy, there should be a threshold for tuning the repressor expression level as extremely low level of repressor may cause high background expression thus reducing the sensing performance. If the desired detection limit cannot be achieved in this way while avoiding high basal expression, other methods can be considered (**Fig. 6.1**), such as moving forward to the second strategy, modifying the intracellular analyte concentration (**Fig. 1.4e**) or

selecting/evolving a repressor with high affinity to its ligand (Li et al. 2015). Also, there should be a threshold for tuning repressor-activators (e.g., MerR) as described in Section 5.2.1. In addition, the activator based sensor should be tuned in the opposite way to improve sensitivity (**Fig. 1.4d**, Wang et al. 2015).

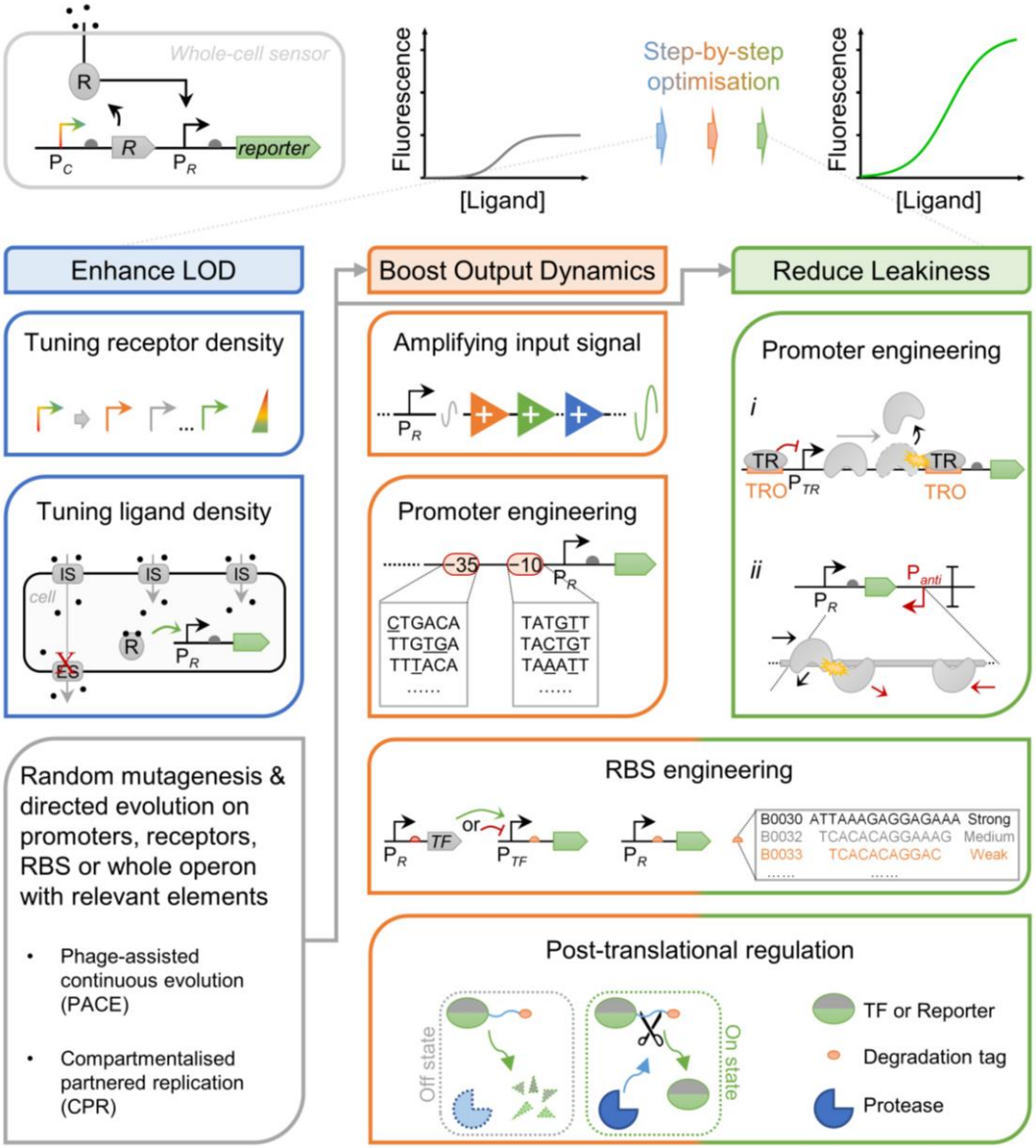


Figure 6.1: Optimisation methodology for whole-cell biosensors.

Whole-cell biosensors can be optimised following a step-by-step methodology: first enhance the sensors' limit of detection (LOD), then boost their output dynamics and finally reduce their leakiness if necessary. Details of each synthetic biology-based optimisation strategy are described in Section 1.3 and Chapter 3. P_C, constitutive promoter. R, receptor. P_R, R's cognate promoter. IS, import system. ES, export system. TF, transcription factor. P_{TF}, TF's cognate promoter. TR, transcriptional repressor. P_{TR}, TR's cognate promoter. TRO, TR's binding site in P_{TR}. P_{anti}, antisense promoter. RBS, ribosome binding site.

For the second strategy, the amplifier optimisation can be adapted to many other transcriptional amplifiers, and the developed amplifier libraries can be applied to many other orthogonal sensors. Although the multi-layered amplifier cascades have been demonstrated successfully, the current amplifier combinations may not be optimal. However, they demonstrate proof of concept that multi-layered amplifier cascades can sequentially amplify transcriptional signals, leading to better output readout and maybe also improving sensor detection limit. The reason that the amplifiers could amplify weak signals and improve the detection limit may be due to the stochastic nature of transcription. It is known that transcription events occur in bursts (Choi et al. 2008), which can be captured by the downstream ultrasensitive transcriptional amplifiers. Therefore, even at low level of input sensor induction, the transduced transcription bursts can still be amplified though at lower frequency, leading to detectable output reporter expression and thus lowering of the sensor detection limit (see Section 3.5). Additionally, many other amplifiers could be identified from other ultrasensitive transcriptional activators to expand the current amplifier library. Notably, it is possible to generate an amplifier data base like the 'Cello' software (Nielsen et al. 2016) for guiding complex genetic circuit design, including characterised input/output threshold, toxicity scores and mathematical modelling, so that researchers can easily find optimal amplifier combinations for their sensors.

Two approaches were tested and combined to reduce arsenic sensor basal expression: promoter structure engineering by adding an extra ArsR binding site (ABS) downstream of the leaky promoter, and protease-based post-translational regulation to regulate the reporter protein degradation.

The first approach alone efficiently reduced sensor leakiness while improving output dynamic range, and it can be applied to any repressor-based sensing system. Moreover, the leakiness was further reduced by increasing the distance between the two ABSs; however, the sensitivity was also decreased. These observations are similar to a previous study (Merulla and van der Meer 2016). This suggests that these effects may be due to the different read-through fraction of RNA polymerase (RNAP) caused by different spacer distances (Hao et al. 2014): at longer distance, the ability of RNAP to dislodge ArsR is lower, so the transcription is blocked better which further reduces output expression at no or low arsenic induction.

The second approach alone can be applied to both activator and repressor-based sensing systems, and it can maintain maximum output expression and sensitivity. However, two limitations were noted (data not shown): **1)** If the protease is located upstream of a reporter gene, the protease DNA sequence may need to be carefully checked to eliminate potential

RNAP binding sequences to avoid increasing output basal expression. This is because heterologous genes from viruses may have low GC content and tend to have multifunctional sequences (thus may contain internal promoters). **2)** This method is not suitable for highly leaky promoters, otherwise the high basal level of protease expression will fully rescue the reduced output and bring back the leakiness. Fortunately, combining the two approaches can sufficiently reduce sensor leakiness while maintaining sensor detection limit and maximum output.

Based on this methodology, I also developed easy-to-interpret sensing platforms for cost effective and portable field testing, where analytes can be readily quantified by simple visualisation. The platforms are based on agarose gel entrapment or microfluidic encapsulation. Both platforms can detect 0.1 ppb – 100/500 ppb arsenite. The former can also detect ≤ 1 ppb arsenate and ≤ 3.18 ppb arsenic from groundwater samples. More importantly, our sensor array had similar sensitivities to both arsenite and arsenate, which will ease the field testing as natural water contains both arsenite and arsenate. Additional tests and modification are still required to prepare the platforms for real applications in the future, such as testing the stability of sensors and medium choice, optimising storage and testing conditions, and employing further physical or intrinsic containment for their safe use in the field (Section 1.5.1 and 1.5.2). Moreover, most currently developed sensors give analog responses, which produce intermediate output and may obscure absolute quantification. To aid in decision making, current sensing systems may be combined with an analog-to-digital converter (e.g., integrases, Section 1.4.1.2) in the future (Rubens et al. 2016).

To further minimise biosafety concerns, the whole cell-based arsenic and mercury sensors were adapted into a crude cell extract-based CFS. Although CFS-based sensors have been successfully demonstrated in previous studies, I noticed some difficulties when transferring the sensors from cells to CFS, such as extreme leakiness of the arsenic sensors, inconsistent results from tuning MerR density and dysfunction of HrpRS/RinA-based amplifiers in CFS. I hypothesised that those difficulties may be due to the intrinsic differences between the CFS and the intracellular environment, such as different levels of macromolecular crowding and different concentrations of TX-TL related components. Hence, targeted optimisation for different sensors in CFS is necessary. Optimisation strategies from the aforementioned methodology were tested to improve CFS-based sensors, as well as some tuning strategies developed specifically for CFS, such as manipulating the quantity of DNA and the addition of pre-expressed regulatory proteins. Most of the strategies were successfully demonstrated although still require further tuning (see Section 5.3). Alternatively, Simcells (Rampley et al.

2017) and artificial cells (Tan et al. 2013; Garamella et al. 2016; Adamala et al. 2017; Rampioni et al. 2018) can be used as cell-free sensor chassis, which should provide a similar intracellular environment to whole cells thus minimising the difference between cellular and cell-free sensors. For the paper-based CFS, additional optimisation is required to reduce background colour and increase signal contrast, to improve the design of paper matrix platform, and to generate sensors for the 'traffic light' system on paper (see Section 5.3).

Overall, the herein presented signal amplifying methodology along with the cellular and CFS-based sensing platforms can be widely applicable to many other cell-based based sensors. The optimisation methods tested here can be readily applied to improve the sensing performance of other cell-based and cell-free genetically-encoded sensors (Pardee et al. 2014; Rampley et al. 2017) to meet their detection requirements for a broad range of real world applications. For example, a recently developed cell-based sensor for detecting pesticides (Khatun et al. 2018) can be further improved and applied to the cellular and CFS-based sensing platforms, which may immediately address the urgent need of the sensitive, cheap and portable detection methods for such major contamination in the environment and our food.

Nevertheless, if the current issues of CFS-based sensors can be addressed in the future, the paper-based cell-free sensing platform will likely be dominant as it will be much easier to use compared to the cellular sensor-based platforms in terms of biosafety concerns, portability, stability, time and cost. Before this comes within reach, however, cellular sensor-based platforms should still be a major player for biosensing.

Appendix 1. Tables for Chapter 2

Appendix Table 2.1: List of genetic parts and sequences used in this study.

Underlined sequences indicate –35 and –10, or –24 and –12 promoter regions. Sequences in red are ArsR binding site (ABS), sequences in blue are MerR binding site, and sequences in yellow are HrpRS binding site, and sequence in italic is TEV cleavage site.

Part name	Type and source	DNA sequence (5'– 3')
J109	Constitutive promoter*	TTTACAGCTAGCTCAGTCCTAGGGACTGTGCTAGC
J117	Constitutive promoter (Wang et al. 2015)	TTGACAGCTAGCTCAGTCCTAGGGATTGTGCTAGC
J114	Constitutive promoter (Wang et al. 2015)	TTTATGGCTAGCTCAGTCCTAGGTACAATGCTAGC
J115	Constitutive promoter (Wang et al. 2015)	TTTATAGCTAGCTCAGCCCTTGGTACAATGCTAGC
J105	Constitutive promoter (Wang et al. 2015)	TTTACGGCTAGCTCAGTCCTAGGTACTATGCTAGC
J106	Constitutive promoter (Wang et al. 2015)	TTTACGGCTAGCTCAGTCCTAGGTATAGTGCTAGC
J101	Constitutive promoter (Wang et al. 2014)	TTTACAGCTAGCTCAGTCCTTGGTATTATGCTAGC
P_{arsR}	Inducible promoter (Wang et al. 2014)	CCAACTCAAAATTCACACCTATTACCTTCCTCTGCACTTACACATTCGTTAAGTCATATATGTTTTTGACTTATCCGCTTCGAAGAGAGACACTACCTGCAA
P_{arsR}-ABS53	Inducible promoter (constructed by PCR)	CCAACTCAAAATTCACACCTATTACCTTCCTCTGCACTTACACATTCGTTAAGTCATATATGTTTTTGACTTATCCGCTTCGAAGAGAGACACTACCTGCAATAAGTCATATATGTTTTTGACTTATCCGCTTC
P_{arsR}-ABS58	Inducible promoter (constructed by PCR)	CCAACTCAAAATTCACACCTATTACCTTCCTCTGCACTTACACATTCGTTAAGTCATATATGTTTTTGACTTATCCGCTTCGAAGAGAGACACTACCTGCAAACTTATAAGTCATATATGTTTTTGACTTATCCGCTTC

P_{arsR}-ABS62	Inducible promoter (constructed by PCR)	CCAACTCAAAATTCACACCTATTACCTTCCTCTGCACTTACACATTCGTT AAGTCATATATGTTTTTGACTTATCCGCTTCGAAGAGAGACACTACCTGC AAACTTACACATAAGTCATATATGTTTTTGACTTATCCGCTTC
P_{arsR}-ABS67	Inducible promoter (constructed by PCR)	CCAACTCAAAATTCACACCTATTACCTTCCTCTGCACTTACACATTCGTT AAGTCATATATGTTTTTGACTTATCCGCTTCGAAGAGAGACACTACCTGC AAACTTACACATTCGTTAAGTCATATATGTTTTTGACTTATCCGCTTC
P_{arsR}-ABS69	Inducible promoter (constructed by PCR)	CCAACTCAAAATTCACACCTATTACCTTCCTCTGCACTTACACATTCGTT AAGTCATATATGTTTTTGACTTATCCGCTTCGAAGAGAGACACTACCTGC AAGCACTTACACATTCGTTAAGTCATATATGTTTTTGACTTATCCGCTTC
P_{arsR}-ABS71	Inducible promoter (constructed by PCR)	CCAACTCAAAATTCACACCTATTACCTTCCTCTGCACTTACACATTCGTT AAGTCATATATGTTTTTGACTTATCCGCTTCGAAGAGAGACACTACCTGC AACTGCACTTACACATTCGTTAAGTCATATATGTTTTTGACTTATCCGCT TC
P_{arsR}-ABS77	Inducible promoter (constructed by PCR)	CCAACTCAAAATTCACACCTATTACCTTCCTCTGCACTTACACATTCGTT AAGTCATATATGTTTTTGACTTATCCGCTTCGAAGAGAGACACTACCTGC AACTTCCTCTGCACTTACACATTCGTTAAGTCATATATGTTTTTGACTTA TCCGCTTC
P_{arsR}-ABS79	Inducible promoter (constructed by PCR)	CCAACTCAAAATTCACACCTATTACCTTCCTCTGCACTTACACATTCGTT AAGTCATATATGTTTTTGACTTATCCGCTTCGAAGAGAGACACTACCTGC AAACCTTCCTCTGCACTTACACATTCGTTAAGTCATATATGTTTTTGACT TATCCGCTTC
P_{arsR}-ABS81	Inducible promoter (constructed by PCR)	CCAACTCAAAATTCACACCTATTACCTTCCTCTGCACTTACACATTCGTT AAGTCATATATGTTTTTGACTTATCCGCTTCGAAGAGAGACACTACCTGC AATTACCTTCCTCTGCACTTACACATTCGTTAAGTCATATATGTTTTTGA CTTATCCGCTTC
P_{arsR}-ABS83	Inducible promoter (constructed by PCR)	CCAACTCAAAATTCACACCTATTACCTTCCTCTGCACTTACACATTCGTT AAGTCATATATGTTTTTGACTTATCCGCTTCGAAGAGAGACACTACCTGC AATATTACCTTCCTCTGCACTTACACATTCGTTAAGTCATATATGTTTTT GACTTATCCGCTTC
P_{arsR}-ABS84	Inducible promoter (constructed by PCR)	CCAACTCAAAATTCACACCTATTACCTTCCTCTGCACTTACACATTCGTT AAGTCATATATGTTTTTGACTTATCCGCTTCGAAGAGAGACACTACCTGC AACTATTACCTTCCTCTGCACTTACACATTCGTTAAGTCATATATGTTTT TGACTTATCCGCTTC
P_{merT}	Inducible promoter (Misra et al. 1985) (*de novo* synthesised)	TTCCATATCGCTTGACTACGTACATGAGTACGGAAGTAAGGTTACGCTAT CCAATCC
P_{hrpL}	Inducible promoter (Wang et al. 2014)	GCCGGATTATGTCCGCTGAGTGGGTCACGGTCCCGGATCAGTTCCCTTGC GAAGCTGACCGATGTTTTTGTGCCAAAAGCTGTTGTGGCAAAAAACGGTT TGCGCAAAGTTTTGTATTACAAAGAATTTCACATTTTAAAATATCTTTAT AAATCAATCAGTTATTTCTATTTTTAAGCTGGCATGGTTATCGCTATAGG GCTTGTAC
Optimised P_{hrpL}	Inducible promoter (constructed by PCR)	GCCGGATTATGTCCGCTGAGTGGGTCACGGTCCCGGATCAGTTCCCTTGC GAAGCTGACCGATGTTTTTGTGCCAAAAGCTGTTGTGGCAAAAAACGGTT TGCGCAAAGTTTTGTATTACAAAGAATTTCACATTTTAAAATATCTTTAT AAATCAATCAGTTATTTCTATTTTCAAGCTGGCACGGTTATTGCTATAGG GCTTGTAC
P_{ecf11_3726}	Inducible promoter (Rhodius et al.	GCCTCCACACCGCTCGTCACATCCTGTGATCCACTCTTCATCCCGCTACG TAACACCTCT

	2013) (*de novo* synthesised)	
P_{ecf16_3622}	Inducible promoter (Rhodius et al. 2013) (*de novo* synthesised)	CTTGGATGAAAAGAAACCCACCGACGGTGTAACCCTGGCGGCCGATGCAACGAACTAACT
P_{ecf17_up1691}	Inducible promoter (Rhodius et al. 2013) (*de novo* synthesised)	CAGTACAAAATTTTTTAGATGCGTTTGGTGAACCAAACTCTTACTCGACTCGTGTCAGTA
P_{ecf20_992}	Inducible promoter (Rhodius et al. 2013) (*de novo* synthesised)	GCGCGGATAAAAATTTCATTTGCCCGCGACGGATTCCCCGCCCATCTATCGTTGAACCCA
P_{ecf22_up1147}	Inducible promoter (Rhodius et al. 2013) (*de novo* synthesised)	CAGTACAAAATTTTTTAGATGCGTTGTTGTGAGGAATCGCGCTCCTGCGCGAATCATCCC
$P_{rinA\_p80\alpha}$	Inducible promoter (Quiles-Puchalt et al. 2013) (*de novo* synthesised)	AATTGGCAGTAAAGTGGCAGTTTTTGATACCTAAAATGAGATATTATGATAGTGTAGGATATTGACTATCTTACTGCGTTTCCCTTATCGCAATTAGGAATAAAGGATCTATGTGGGTTGGCTGATTATAGCCAATCCTTTTTTAATTTTAAAAAGCGTATAGCGCGAG
B0030	RBS**	ATTAAAGAGGAGAAA
B0031	RBS**	TCACACAGGAAACC
B0032	RBS**	TCACACAGGAAAG
B0033	RBS**	TCACACAGGAC
B0034	RBS**	AAAGAGGAGAAA
B0015	Terminator***	CCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATA
ASV	Degradation tag (Andersen et al. 1998)	GCTGCAAACGACGAAAACTACGCTGCATCAGTT
AAV	Degradation tag (Andersen et al. 1998)	GCTGCAAACGACGAAAACTACGCTGCCGCAGTT
te	TEV cleavage site (Fernandez-Rodriguez and Voigt 2016) with linker	GGCGGCTCCGGCGGCTCCGGCGGCACCGGCGGCTCCGGCGGCTCC*GAAAA CCTGTATTTTCAG*GGTGGC

gfp	Gene (Wang et al. 2014)	ATGCGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGA ATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTG AAGGTGATGCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACT GGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCGGTTATGG TGTTCAATGCTTTGCGAGATACCCAGATCATATGAAACAGCATGACTTTT TCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTC AAAGATGACGGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGA TACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATG GAAACATTCTTGGACACAAATTGGAATACAACTATAACTCACACAATGTA TACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAAT TAGACACAACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAAC AAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTAC CTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAGAGAGACCA CATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACATGGCATGG ATGAACTATACAAATAA
mRFP	Gene****	ATGGCTTCCTCCGAAGACGTTATCAAAGAGTTCATGCGTTTCAAAGTTCG TATGGAAGGTTCCGTTAACGGTCACGAGTTCGAAATCGAAGGTGAAGGTG AAGGTCGTCCGTACGAAGGTACCCAGACCGCTAAACTGAAAGTTACCAAA GGTGGTCCGCTGCCGTTCGCTTGGGACATCCTGTCCCCGCAGTTCCAGTA CGGTTCCAAAGCTTACGTTAAACACCCGGCTGACATCCCGGACTACCTGA AACTGTCCTTCCCGGAAGGTTTCAAATGGGAACGTGTTATGAACTTCGAA GACGGTGGTGTTGTTACCGTTACCCAGGACTCCTCCCTGCAAGACGGTGA GTTCATCTACAAAGTTAAACTGCGTGGTACCAACTTCCCGTCCGACGGTC CGGTTATGCAGAAAAAAACCATGGGTTGGGAAGCTTCCACCGAACGTATG TACCCGGAAGACGGTGCTCTGAAAGGTGAAATCAAAATGCGTCTGAAACT GAAAGACGGTGGTCACTACGACGCTGAAGTTAAAACCACCTACATGGCTA AAAAACCGGTTCAGCTGCCGGGTGCTTACAAAACCGACATCAAACTGGAC ATCACCTCCCACAACGAAGACTACACCATCGTTGAACAGTACGAACGTGC TGAAGGTCGTCACTCCACCGGTGCTTAA
mCherry	Gene*****	ATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCAT GCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGA TCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAG CTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTC CCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACA TCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGC GTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTC CTTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACT TCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCC TCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAA GCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGA CCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTC AACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGA ACAGTACGAACGCGCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGC TGTACAAGTAA
lacZ	Gene (Amplified from *E. coli* MG1655 by PCR)	ATGACCATGATTACGGATTCACTGGCCGTCGTTTTACAACGTCGTGACTG GGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTT TCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAA CAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCACC AGAAGCGGTGCCGGAAAGCTGGCTGGAGTGCGATCTTCCTGAGGCCGATA CTGTCGTCGTCCCCTCAAACTGGCAGATGCACGGTTACGATGCGCCCATC TACACCAACGTGACCTATCCCATTACGGTCAATCCGCCGTTTGTTCCCAC GGAGAATCCGACGGGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCT GGCTACAGGAAGGCCAGACGCGAATTATTTTTGATGGCGTTAACTCGGCG TTTCATCTGTGGTGCAACGGGCGCTGGGTCGGTTACGGCCAGGACAGTCG TTTGCCGTCTGAATTTGACCTGAGCGCATTTTTACGCGCCGGAGAAAACC GCCTCGCGGTGATGGTGCTGCGCTGGAGTGACGGCAGTTATCTGGAAGAT CAGGATATGTGGCGGATGAGCGGCATTTTCCGTGACGTCTCGTTGCTGCA

		TAAACCGACTACACAAATCAGCGATTTCCATGTTGCCACTCGCTTTAATG ATGATTTCAGCCGCGCTGTACTGGAGGCTGAAGTTCAGATGTGCGGCGAG TTGCGTGACTACCTACGGGTAACAGTTTCTTTATGGCAGGGTGAAACGCA GGTCGCCAGCGGCACCGCGCCTTTCGGCGGTGAAATTATCGATGAGCGTG GTGGTTATGCCGATCGCGTCACACTACGTCTGAACGTCGAAAACCCGAAA CTGTGGAGCGCCGAAATCCCGAATCTCTATCGTGCGGTGGTTGAACTGCA CACCGCCGACGGCACGCTGATTGAAGCAGAAGCCTGCGATGTCGGTTTCC GCGAGGTGCGGATTGAAAATGGTCTGCTGCTGCTGAACGGCAAGCCGTTG CTGATTCGAGGCGTTAACCGTCACGAGCATCATCCTCTGCATGGTCAGGT CATGGATGAGCAGACGATGGTGCAGGATATCCTGCTGATGAAGCAGAACA ACTTTAACGCCGTGCGCTGTTCGCATTATCCGAACCATCCGCTGTGGTAC ACGCTGTGCGACCGCTACGGCCTGTATGTGGTGGATGAAGCCAATATTGA AACCCACGGCATGGTGCCAATGAATCGTCTGACCGATGATCCGCGCTGGC TACCGGCGATGAGCGAACGCGTAACGCGAATGGTGCAGCGCGATCGTAAT CACCCGAGTGTGATCATCTGGTCGCTGGGGAATGAATCAGGCCACGGCGC TAATCACGACGCGCTGTATCGCTGGATCAAATCTGTCGATCCTTCCCGCC CGGTGCAGTATGAAGGCGGCGGAGCCGACACCACGGCCACCGATATTATT TGCCCGATGTACGCGCGCGTGGATGAAGACCAGCCCTTCCCGGCTGTGCC GAAATGGTCCATCAAAAAATGGCTTTCGCTACCTGGAGAGACGCGCCCGC TGATCCTTTGCGAATACGCCCACGCGATGGGTAACAGTCTTGGCGGTTTC GCTAAATACTGGCAGGCGTTTCGTCAGTATCCCCGTTTACAGGGCGGCTT CGTCTGGGACTGGGTGGATCAGTCGCTGATTAAATATGATGAAAACGGCA ACCCGTGGTCGGCTTACGGCGGTGATTTTGGCGATACGCCGAACGATCGC CAGTTCTGTATGAACGGTCTGGTCTTTGCCGACCGCACGCCGCATCCAGC GCTGACGGAAGCAAAACACCAGCAGCAGTTTTTCCAGTTCCGTTTATCCG GGCAAACCATCGAAGTGACCAGCGAATACCTGTTCCGTCATAGCGATAAC GAGCTCCTGCACTGGATGGTGGCGCTGGATGGTAAGCCGCTGGCAAGCGG TGAAGTGCCTCTGGATGTCGCTCCACAAGGTAAACAGTTGATTGAACTGC CTGAACTACCGCAGCCGGAGAGCGCCGGGCAACTCTGGCTCACAGTACGC GTAGTGCAACCGAACGCGACCGCATGGTCAGAAGCCGGGCACATCAGCGC CTGGCAGCAGTGGCGTCTGGCGGAAAACCTCAGTGTGACGCTCCCCGCCG CGTCCCACGCCATCCCGCATCTGACCACCAGCGAAATGGATTTTTGCATC GAGCTGGGTAATAAGCGTTGGCAATTTAACCGCCAGTCAGGCTTTCTTTC ACAGATGTGGATTGGCGATAAAAAACAACTGCTGACGCCGCTGCGCGATC AGTTCACCCGTGCACCGCTGGATAACGACATTGGCGTAAGTGAAGCGACC CGCATTGACCCTAACGCCTGGGTCGAACGCTGGAAGGCGGCGGGCCATTA CCAGGCCGAAGCAGCGTTGTTGCAGTGCACGGCAGATACACTTGCTGATG CGGTGCTGATTACGACCGCTCACGCGTGGCAGCATCAGGGGAAAACCTTA TTTATCAGCCGGAAAACCTACCGGATTGATGGTAGTGGTCAAATGGCGAT TACCGTTGATGTTGAAGTGGCGAGCGATACACCGCATCCGGCGCGGATTG GCCTGAACTGCCAGCTGGCGCAGGTAGCAGAGCGGGTAAACTGGCTCGGA TTAGGGCCGCAAGAAAACTATCCCGACCGCCTTACTGCCGCCTGTTTTGA CCGCTGGGATCTGCCATTGTCAGACATGTATACCCCGTACGTCTTCCCGA GCGAAAACGGTCTGCGCTGCGGGACGCGCGAATTGAATTATGGCCCACAC CAGTGGCGCGGCGACTTCCAGTTCAACATCAGCCGCTACAGTCAACAGCA ACTGATGGAAACCAGCCATCGCCATCTGCTGCACGCGGAAGAAGGCACAT GGCTGAATATCGACGGTTTCCATATGGGGATTGGTGGCGACGACTCCTGG AGCCCGTCAGTATCGGCGGAGTTTCAGCTGAGCGCCGGTCGCTACCATTA CCAGTTGGTCTGGTGTCAAAAATAA
arsR	Gene (Wang et al. 2014)	ATGTCATTTCTGTTACCCATCCAATTGTTCAAAATTCTTGCTGATGAAAC CCGTCTGGGCATCGTTTTACTGCTCAGCGAACTGGGAGAGTTATGCGTCT GCGATCTCTGCACTGCTCTCGACCAGTCGCAGCCCAAGATCTCCCGCCAC CTGGCATTGCTGCGTGAAAGCGGGCTATTGCTGGACCGCAAGCAAGGTAA GTGGGTTCATTACCGCTTATCACCGCATATTCCAGCATGGGCGGCGAAAA TTATTGATGAGGCCTGGCGATGTGAACAGGAAAAGGTTCAGGCGATTGTC CGCAACCTGGCTCGACAAAACTGTTCCGGGGACAGTAAGAACATTTGCAG TTAA

merR	Gene (Misra et al. 1985) (*de novo* synthesised)	ATGGAAAATAATTTGGAAAACCTGACCATTGGCGTTTTTGCCAAGGCGGC CGGGGTCAACGTGGAGACAATCCGCTTCTATCAGCGCAAGGGCCTGTTGC GGGAACCGGACAAGCCTTACGGCAGCATCCGCCGCTATGGGGAGGCGGAC GTGGTTCGGGTGAAATTCGTGAAATCGGCACAGCGGCTGGGGTTCAGTCT GGACGAGATTGCCGAGCTGTTGCGGCTCGACGATGGCACCCACTGCGAGG AGGCCAGCAGCCTGGCCGAACACAAGCTCAAGGACGTGCGCGAGAAGATG GCCGACTTGGCGCGCATGGAAACCGTGCTGTCTGAACTCGTGTGCGCCTG CCATGCACGAAAGGGGAATGTTTCCTGCCCGTTGATCGCGTCACTACAGG GCGAAGCAGGCCTGGCAAGGTCAGCTATGCCTTAG
hrpR	Gene (Wang et al. 2014)	ATGAGTACAGGCATCGATAAGGACGTCCGAGAGTGTTGGGGCGTAACTGC ATTATCAGCGGGTCATCAAATTGCAATGAATAGCGCGTTTCTGGATATGG ACTTGCTGTTGTGCGGGGAAACCGGCACCGGCAAGGACACACTGGCCAAC CGCATTCACGAGTTGTCCAGCAGGTCGGGACCCTTTGTGGGCATGAACTG CGCCGCCATTCCCGAGTCGCTGGCAGAGAGCCAGTTATTCGGTGTGGTCA ACGGTGCATTCACCGGCGTATGCCGGGCTCGCGAGGGCTACATAGAGGCC TCCAGTGGTGGCACCTTGTACCTGGATGAAATCGACAGCATGCCGTTGAG CCTGCAAGCCAAACTGCTGCGTGTGTTGGAGAGTCGAGGTATCGAGCGTC TGGGCTCGACCGAATTTATCCCGGTGGATCTGCGGATCATTGCCTCGGCC CAGCGGCCACTGGATGAACTGGTGGAACAAGGACTTTTCCGTCGCGACCT GTTTTTTCGGCTCAACGTGCTGACGCTTCACTTGCCAGCCTTGCGCAAAC GTCGTGAACAGATCCTGCCATTGTTCGACCAGTTCACCCAGGGTATCGCT GCCGAGTTCGGACGTCCCGCTCCTGCGCTGGACAGCGGGCGTGTGCAGCT GCTGCTCAGCCACGACTGGCCGGGCAACATCCGCGAATTGAAGTCTGCGG CCAAGCGCTTCGTACTCGGCTTCCCCTTGCTGGGCGCCGACCCTGTGGAA GCGCTTGACCCTGCCACGGGGCTGCGCACGCAAATGCGCATCATCGAGAA AATGCTCATCCAGGATGCCTTGAAGCGGCACAGGCACAATTTCGACGCGG TGCTTCAGGAGTTGGAGTTGCCAAGACGCACCCTGTATCACCGCATGAAG GAACTGGGAGTTGCAGCGCCGATCGCTGCGACGGCCGGGGTCTAA
hrpS	Gene (Wang et al. 2014)	ATGAGTCTTGATGAAAGGTTTGAGGATGATCTGGACGAGGAGCGGGTTCC GAATCTGGGGATAGTTGCCGAAAGTATTTCGCAACTGGGTATCGACGTGC TGCTATCGGGTGAGACCGGCACGGGCAAAGACACGATTGCCCGACGGATT CATGAGATGTCAGGCCGCAAAGGGCGCCTGGTGGCGATGAATTGCGCGGC CATTCCGGAGTCCCTCGCCGAGAGCGAGTTATTCGGCGTGGTCAGCGGTG CCTACACCGGCGCTGATCGCTCCAGAGTCGGTTATGTCGAAGCGGCGCAG GGCGGCACGCTGTACCTGGATGAGATCGATAGCATGCCGCTGAGCCTGCA AGCCAAATTGCTGAGGGTGCTGGAAACCCGAGCGCTTGAACGGCTGGGTT CGACGTCGACGATCAAGCTGGATATCTGCGTGATCGCCTCCGCCCAATGC TCGCTGGACGACGCCGTCGAGCGGGGGCAGTTTCGTCGCGATCTGTATTT TCGCCTGAACGTCCTGACACTCAAGCTTCCTCCGCTACGTAACCAGTCTG ATCGCATAGTTCCCCTGTTCACACGTTTTACGGCCGCCGCCGCGAGGGAG CTCGGTGTTCCCGTTCCCGATGTTTGCCCACTGCTGCACAAAGTGCTGCT GGGCCACGACTGGCCCGGCAATATCCGTGAGCTCAAGGCGGCAGCCAAAC GCCATGTGCTGGGTTTCCCCTTGCTGGGCGCCGAGCCGCAGGGCGAAGAG CACTTGGCCTGTGGGCTCAAATCGCAATTGCGAGTGATCGAAAAAGCCCT GATTCAGGAGTCGCTCAAGCGCCACGACAATTGTGTGGATTCGGTAAGCC TGGAACTGGACGTGCCACGCCGTACGCTCTATCGACGCATCAAAGAATTG CAGATCTAA

ECF11_987	Gene (Rhodius et al. 2013) (*de novo* synthesised)	ATGATGAGCGATAGTCCGCAGAAACTGGGTCGTAATGAATGGAATGCATA TATGGATAAAGTGAAAGCCAAAGATCGTGAAGCCTTTGCCTTTGTGTTTC GTTTTTATGCACCGAAACTGAAACAGTTCGCCTATAAACATGTTGGCAAT GAACAGGTTGCAATGGAAATGGTTCAAGAAACCATGGCAACCGTTTGGCA GAAAGCACATCTGTATGATGGTAAAAAAAGCGCACTGAGCACCTGGATTT ATACCATTATTCGTAACCTGTGCTTTGACCTGCTGCGTAAACAGAAAGGT AAAGAACTGCATATTCACAGCGACGATATTTGGCCGAGCGAATATTATCC GCCTGATATGGTTGATCATTATAGTCCGGAACAGGATATGCTGAAAGAAC AGGTGGTTAAATTTCTGGATATCCTGCCGAAAAATCAGCGTGATGTTCTG CAAGCAGTTTATCTGGAAGAACTGCCGCATCAGCAGGTTGCAGAACTGTT TGATATTCCGCTGGGCACCGTTAAAAGCCGTCTGCGTCTGGCAGTTGAAA AACTGCGTCATAGCATGCATACCGAACAGCTGTAA
ECF16_3622	Gene (Rhodius et al. 2013) (Addgene #49670 with modification)	ATGCAGCGTACCAATAGCCAGGATGTTCTGAGCACCCGTGAAAGCCAGCT GCAAGCATTACTGCTGCAAGGTCTGGCAGGCGATACCTTTGCATATCGTC AGTTTCTGACCGCACTGGCAGCACATATTCGTGGTTTTCTGCGTCGTCGT CTGCCGCAGCATCCGGCAGAAGTTGAAGATCTGCTGCAAGAAGTTCTGCT GGCAGTTCATAATGCACGTCATACCTATCAGGCACGTCAGCCGCTGACCG CATGGGTTCAGGCAATTGCACGTTATAAACTGGCAGATCATCTGCGTAGC CATGCACGTCGTGAAGCACGTCATGATCTGCTGGATGATGATAGCGAACT GTTTGCAGCAAGTGATGAACAGCCTGCACAGGCAAGCCGTGATCTGGGTA AACTGCTGGGTCAGCTGCCGGATCGTCAGCGTCTGCCGATTGTTCATGTT AAACTGGAAGGTCTGAGCGTTGAAGAAACCGCACAGATTACCGGTCTGAG CAGCAGCGCAGTTAAAGTTGGTATTCATCGTGGTCTGAAAGCACTGGGCA AACTGATTCGTGGTAAAGGTCATGATGAAGATCGTTAA
ECF17_1691	Gene (Rhodius et al. 2013) (Addgene #49672)	ATGGCACGTGTTAGCGGTGCAGCAGCAGCCGAAGCAGCACTGATGCGTGC ACTGTATGATGAACATGCAGCAGTTCTGTGGCGTTATGCACTGCGTCTGA CCGGTGATGCAGCACAGGCAGAAGATGTTGTTCAAGAAACCCTGCTGCGT GCATGGCAGCATCCGGAAGTTATTGGTGATACCGCACGTCCGGCACGCGC ATGGCTGTTTACCGTTGCACGTAATATGATTATTGATGAACGTCGTAGCG CACGTTTTCGTAATGTTGTTGGTAGCACCGATCAGAGCGGTACACCGGAA CAGAGCACACCGGATGAAGTTAATGCAGCACTGGATCGTCTGCTGATTGC AGATGCACTGGCACAGCTGAGCGCAGAACATCGTGCAGTTATTCAGCGTA GCTATTATCGTGGTTGGAGCACCGCACAGATTGCAACCGATCTGGGTATT GCAGAAGGCACCGTTAAAAGCCGTCTGCATTATGCAGTTCGTGCCCTGCG CCTGACCCTGCAAGAACTGGGTGTTACCCGTTAA
ECF20_992	Gene (Rhodius et al. 2013) (Addgene #49673 with modification)	ATGAATGAAACCGATCCTGATCTGGAACTGCTGAAACGTATTGGTAATAA TGATGCACAGGCCGTTAAAGAAATGGTTACCCGTAAACTGCCTCGTCTGC TGGCACTGGCAAGTCGCCTGCTGGGTGATGCAGATGAAGCACGTGATATT GCACAAGAAAGTTTTCTGCGCATTTGGAAACAGGCAGCAAGCTGGCGTAG CGAACAGGCACGTTTTGATACCTGGCTGCATCGTGTTGCACTGAATCTGT GTTATGATCGTCTGCGTCGTCGTAAAGAACATGTGCCGGTTGATAGCGAA CATGCCTGTGAAGCACTGGATACCCGTCCGGCACCGGATGAACAGCTGGA AGCAAGCGCACAGAGCCGTCGTATGGCACAGGCACTGGATCAGCTGCCGG ATCGTCAGCGTGAAGCAATTGTTCTGCAATATTATCAAGAACTGAGCAAT ACCGAAGCAGCAGCACTGATGCAAATTAGCGTTGAAGCCCTGGAAAGCCT GCTGAGCCGTGCACGTCGTAATCTGCGTAGCCATCTGGCCGAAGCACCGG GTGCAGATCTGAGCGGTCGTCGCAAACCGTAA

ECF22_4450	Gene (Rhodius et al. 2013) (Addgene #49676)	ATGCCGCAGCAGACCGATCTGCATGCACGTTTTAGCGCACTGATGCAGCA GCATCGTGGTATTGTTCTGAAAGTTGCAGCAAGCTATTGTCGTGATCCGG ATGATCGTGCAGATCTGGCACAGGATATTGCAACCCATCTGTGGCGTGCA TTTCCGAGCTATGATCCGCATCGTCGTTTTAGCACCTGGATGTATCGTAT TGCACTGAATGTTGCAATTAGCGATCTGCGTAGCCGTCGTCATGCACCGA GCGAAGTTCTGGATGTTGGCACCCTGGTTGATAGCATTGCAGATCCGCAT GCTCGCGATCCGGCACGTGAACATCAGCTGACCGCACTGTATACCTTTAT TGCACAGCTGCCTCGTCTGGAACGTGCCCTGATGCTGCTGTATCTGGATG ATCATAGTTATCGTGAAATTGCAGATGTGCTGGGTATTAGCGAAACCAAT GTTGCCACCAAACTGAGCCGTCTGAAAGCACGTATTCGTGCAGAACTGTA ATAA
rinA_p80α	Gene (Quiles-Puchalt et al. 2013) (*de novo* synthesised)	ATGACTAAAAAGAAATATGGATTAAAATTATCAACAGTTCGAAAGTTAGA AGATGAGTTGTGTGATTATCCTAATTATCATAAGCAACTCGAAGATTTAA GAAGTGAAATAATGACACCATGGATTCCAACAGATACAAATATAGGCGGG GAGTTTGTACCGTCTAATACATCGAAAACAGAAATGGCAGTAACTAATTA TCTTTGTAGTATACGAAGAGGTAAAATCCTTGAGTTTAAGAGCGCTATTG AACGTATAATCAACACATCAAGTAGGAAAGAACGCGAATTTATTCAAGAG TATTATTTTAATAAAAAGGAATTAGTGAAAGTTTGTGATGACATACACAT TTCTGATAGAACTGCTCATAGAATCAAAAGGAAAATCATATCCAGATTGG CGGAAGAGTTAGGGGAAGAGTGA
malE	Gene (Amplified from *E. coli* MG1655 by PCR)	ATGATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAGGCTA TAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTA AAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTT GCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTT TGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAG CGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAAC GGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTA TAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGG CGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC CTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTA TGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAAGACGTGGGCGTGG ATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTAAA AACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTT TAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCA ACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTC AAGGGTCAACCATCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAA CGCCGCCAGTCCGAACAAAGAGCTGGCGAAAGAGTTCCTCGAAAACTATC TGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAACCGCTGGGT GCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTAT TGCCGCCACCATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCC CGCAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCC GCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTCG TATCACCAAGTAA

TEV	Gene******	ATGGGCGAAAGCCTGTTCAAAGGTCCGCGTGACTACAACCCGATTAGCTC GACGATCTGCCACCTGACGAACGAAAGCGACGGCCACACCACGAGCCTGT ATGGCATCGGTTTTGGCCCGTTCATTATCACGAACAAACACCTGTTTCGT CGCAACAATGGTACCCTGCTGGTGCAGTCTCTGCATGGCGTGTTTAAAGT TAAAAATACCACGACCCTGCAACAACACCTGATCGATGGTCGTGACATGA TTATCATTCGCATGCCGAAAGATTTTCCGCCGTTCCCGCAGAAACTGAAA TTCCGTGAACCGCAACGTGAAGAACGCATTTGCCTGGTCACGACCAACTT TCAGACCAAATCAATGAGCTCTATGGTTAGCGATACGTCTTGTACCTTTC CGAGTTCCGACGGCATCTTCTGGAAACATTGGATTCAGACCAAAGATGGT CAATGCGGCAGTCCGCTGGTTTCCACCCGTGACGGTTTCATCGTCGGCAT TCACTCAGCGTCGAACTTTACGAATACCAACAATTACTTCACGTCCGTTC CGAAAAACTTTATGGAACTGCTGACCAATCAGGAAGCGCAGCAATGGGTG TCAGGTTGGCGCCTGAATGCCGATTCGGTTCTGTGGGGCGGTCACAAAGT CTTTATGGTGAAACCGGAAGAACCGTTCCAGCCGGTCAAAGAAGCAACCC AACTGATGAATTAA

*: http://parts.igem.org/Promoters/Catalog/Constitutive.

**: http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Community_Collection.

***: http://parts.igem.org/Part:BBa_B0015.

****: http://parts.igem.org/Part:BBa_E1010.

*****: http://parts.igem.org/Part:BBa_J06504.

******: http://parts.igem.org/Part:BBa_K1319004.

Appendix Table 2.2: List of plasmid constructs used in this study.

J/P represents promoter, 30/31/32/33/34 represents RBS, letters in italic represent the name of a gene or a short peptide, t is the terminator B0015, and Amp represents amplifier. P_{hrpL}^{E} is the optimised P_{hrpL}.

Plasmid	Description	Reference
pSB4A3	BioBrick vector, pSC101 ori, Amp^R	Shetty et al. 2008
pSB3K3	BioBrick vector, p15A ori, Kan^R	Shetty et al. 2008
pSB1K3	BioBrick vector, pMB1 ori, Kan^R	Shetty et al. 2008
pSB1A3	BioBrick vector, pMB1 ori, Amp^R	Shetty et al. 2008
pSB1C3	BioBrick vector, pMB1 ori, Cam^R	Shetty et al. 2008
pXWJ109-gfp	pSB3K3 carrying J109-30*gfp*-t	This study
pBW201J117-gfp	pSB3K3 carrying J117-30*gfp*-t	Wang et al. 2015
pBW202J114-gfp	pSB3K3 carrying J114-30*gfp*-t	Wang et al. 2015
pBW204J115-gfp	pSB3K3 carrying J115-30*gfp*-t	Wang et al. 2015
pBW206J106-gfp	pSB3K3 carrying J106-30*gfp*-t	Wang et al. 2015
pXWJ101-gfp	pSB3K3 carrying J101-30*gfp*-t	This study
pXWJ101ArsR	pSB1K3 carrying J101-32*arsR*-t	This study
pXWJ101ArsR2	pSB3K3 carrying J101-32*arsR*-t	This study
pXWJ101ArsR3	pSB4A3 carrying J101-32*arsR*-t	This study
pXWParsR-mCh	pSB1A3 carrying P_{arsR}-34*mCherry*-t	This study
pXWParsR-mCh2	pSB3K3 carrying P_{arsR}-34*mCherry*-t	This study
pXWABS62-mCh	pSB1A3 carrying P_{arsR}-ABS62-34*mCherry*-t	This study
pXWABS67-mCh	pSB1A3 carrying P_{arsR}-ABS67-34*mCherry*-t	This study
pXWABS69-mCh	pSB1A3 carrying P_{arsR}-ABS69-34*mCherry*-t	This study
pXWABS71-mCh	pSB1A3 carrying P_{arsR}-ABS71-34*mCherry*-t	This study
pXWABS77-mCh	pSB1A3 carrying P_{arsR}-ABS77-34*mCherry*-t	This study
pXWABS79-mCh	pSB1A3 carrying P_{arsR}-ABS79-34*mCherry*-t	This study
pXWABS81-mCh	pSB1A3 carrying P_{arsR}-ABS81-34*mCherry*-t	This study
pXWABS83-mCh	pSB1A3 carrying P_{arsR}-ABS83-34*mCherry*-t	This study
pXWABS84-mCh	pSB1A3 carrying P_{arsR}-ABS84-34*mCherry*-t	This study
pXWcABS10-mCh	pSB1A3 carrying cABS10-34*mCherry*-t	This study
pXWdABS10-mCh	pSB1A3 carrying dABS10-34*mCherry*-t	This study
pXWeABS10-mCh	pSB1A3 carrying eABS10-34*mCherry*-t	This study
pXW103ABS10-mCh	pSB1A3 carrying 103ABS10-34*mCherry*-t	This study
pXW105ABS10-mCh	pSB1A3 carrying 105ABS10-34*mCherry*-t	This study
pXWParsR11-mCh	pSB1A3 carrying P_{arsR}11-34*mCherry*-t	This study
pXWParsR12-mCh	pSB1A3 carrying P_{arsR}12-34*mCherry*-t	This study
pXWParsR13-mCh	pSB1A3 carrying P_{arsR}13-34*mCherry*-t	This study

pXWParsR14-mCh	pSB1A3 carrying P_{arsR}14-34*mCherry*-t	This study
pXWParsR11-mCh	pSB1A3 carrying P_{arsR}15-34*mCherry*-t	This study
pXWParsR21-mCh	pSB3K3 carrying P_{arsR}21-34*mCherry*-t	This study
pXWParsR22-mCh	pSB1A3 carrying P_{arsR}22-34*mCherry*-t	This study
pXWParsR24-mCh	pSB1A3 carrying P_{arsR}24-34*mCherry*-t	This study
pXWParsR25-mCh	pSB3K3 carrying P_{arsR}25-34*mCherry*-t	This study
pXWParsR31-mCh	pSB1A3 carrying P_{arsR}31-34*mCherry*-t	This study
pXWParsR32-mCh	pSB1A3 carrying P_{arsR}32-34*mCherry*-t	This study
pXWParsR33-mCh	pSB1A3 carrying P_{arsR}33-34*mCherry*-t	This study
pXWParsR34-mCh	pSB1A3 carrying P_{arsR}34-34*mCherry*-t	This study
pXWParsR35-mCh	pSB1A3 carrying P_{arsR}35-34*mCherry*-t	This study
pXWParsR41-mCh	pSB1A3 carrying P_{arsR}41-34*mCherry*-t	This study
pXWParsR42-mCh	pSB1A3 carrying P_{arsR}42-34*mCherry*-t	This study
pXWParsR43-mCh	pSB3K3 carrying P_{arsR}43-34*mCherry*-t	This study
pXWParsR44-mCh	pSB3K3 carrying P_{arsR}44-34*mCherry*-t	This study
pXWParsR45-mCh	pSB1A3 carrying P_{arsR}45-34*mCherry*-t	This study
pXWParsR51-mCh	pSB1A3 carrying P_{arsR}51-34*mCherry*-t	This study
pXWParsR52-mCh	pSB1A3 carrying P_{arsR}52-34*mCherry*-t	This study
pXWParsR53-mCh	pSB1A3 carrying P_{arsR}53-34*mCherry*-t	This study
pXWParsR54-mCh	pSB1A3 carrying P_{arsR}54-34*mCherry*-t	This study
pXWParsR55-mCh	pSB1A3 carrying P_{arsR}55-34*mCherry*-t	This study
pBW100ParsR	pSB3K3 carrying J101-32*arsR*-t-P_{arsR}	Wang et al. 2014
pXWJ117As	pSB3K3 carrying J117-30*arsR*-t-P_{arsR}	This study
pXWJ109As	pSB3K3 carrying J109-30arsR-t-P_{arsR}	This study
pBW101ParsR-gfp	pSB3K3 carrying J101-32*arsR*-t-P_{arsR}-30*gfp*-t	Wang et al. 2014
pBW411J117-arsR	pSB3K3 carrying J117-30*arsR*-t-P_{arsR}-30*gfp*-t	Wang et al. 2015
pBW412J114-arsR	pSB3K3 carrying J114-30*arsR*-t-P_{arsR}-30*gfp*-t	Wang et al. 2015
pBW414J105-arsR	pSB3K3 carrying J105-30*arsR*-t-P_{arsR}-30*gfp*-t	Wang et al. 2015
pXWJ109As-gfp	pXWJ109As encoding 30*gfp*-t	This study
pBW103ParsR-Amp30^{C}	pBW100ParsR carrying 30*hrpR*-30*hrpS*-t-P_{hrpL}-30*gfp*-t	Wang et al. 2014
pXWJ109AsAmp30^{C}	pXWJ109As encoding 30*hrpR*-30*hrpS*-t-P_{hrpL}-30*gfp*-t	This study
pXWJ109AsAmp30^{E}	pXWJ109As encoding 30*hrpR*-30*hrpS*-t-P_{hrpL}^{E}-30*gfp*-t	This study
pXWJ117AsAmp30^{E}	pXWJ117As encoding 30*hrpR*-30*hrpS*-t-P_{hrpL}^{E}-30*gfp*-t	This study
pXWJ117AsAmp30^{E}2	pSB4A3 carrying J117-30*arsR*-t-P_{arsR}-30*hrpR*-30*hrpS*-t-P_{hrpL}^{E}-30*gfp*-t	This study
pXWJ115MerR	pSB3K3 carrying J115-32*merR*-t	This study
pXWPmerT-LacZ	pSB1A3 carrying P_{merT}-30*lacZ*-t	This study
pXWJ115Hg	pSB3K3 carrying J115-32*merR*-t-P_{merT}	This study
pXWJ117Hg	pSB3K3 carrying J117-32*merR*-t-P_{merT}	This study

pXWJ114Hg	pSB3K3 carrying J114-32*merR*-t-P_{merT}	This study
pXWJ109Hg	pSB3K3 carrying J109-32*merR*-t-P_{merT}	This study
PmerT-rbs30-gfp	pXWJ115Hg encoding 30*gfp*-t	Wang and Buck 2014
pXWJ114Hg-gfp	pXWJ114Hg encoding 30*gfp*-t	This study
pXWJ109Hg-gfp	pXWJ109Hg encoding 30*gfp*-t	This study
pXWJ115Hg-mCh	pXWJ115Hg encoding 34*mCherry*-t	This study
pXWJ117Hg-mCh	pXWJ117Hg encoding 34*mCherry*-t	This study
pXWJ109Hg-mCh	pXWJ109Hg encoding 34*mCherry*-t	This study
pXWJ115Hg-LacZ	pXWJ115Hg encoding 30*lacZ*-t	This study
pXWJ109HgAmp30^{E}	pXWJ109Hg encoding 30*hrpR*-30*hrpS*-t-P_{hrpL}^{E}-30*gfp*-t	This study
pXWAmp30^{E}	pSB3K3 carrying Amp30^{E} which is 30*hrpR*-30*hrpS*-t-P_{hrpL}^{E}	This study
pXWJ109HgRS	pXWJ109Hg encoding 30*hrpR*-30*hrpS*-t	This study
pBW400hrpL-gfp	pSB4A3 carrying P_{hrpL}-30*gfp*-t	Wang et al. 2011
pXWPhrpL-gfp	pSB4A3 carrying P_{hrpL}^{E}-30*gfp*-t	This study
pXWJ109HgRinA	pXWJ109Hg encoding 30*rinA*-t	This study
pXWPrinA-gfp1	pSB4A3 carrying P_{rinA}-30*gfp*-t	This study
pXWJ109HgE11	pXWJ109Hg encoding 30*E11*-t	This study
pXWP11-gfp1	pSB4A3 carrying P_{e11}-30*gfp*-t	This study
pXWAmp30R	pSB3K3 carrying Amp30R which is 30*RinA*-t-P_{rinA}	This study
pXWAmp30R^{A}	pSB3K3 carrying Amp30R^{A} which is 30*rinA-ASV*-t-P_{rinA}	This study
pXWAmp31R	pSB3K3 carrying Amp31R which is 31*rinA*-t-P_{rinA}	This study
pXWAmp31R^{A}	pSB3K3 carrying Amp31R^{A} which is 31*rinA-ASV*-t-P_{rinA}	This study
pXWAmp32R	pSB3K3 carrying Amp32R which is 32*rinA*-t-P_{rinA}	This study
pXWAmp32R^{A}	pSB3K3 carrying Amp32R^{A} which is 32*rinA-ASV*-t-P_{rinA}	This study
pXWAmp33R	pSB3K3 carrying Amp33R which is 33*rinA*-t-P_{rinA}	This study
pXWAmp33R^{A}	pSB3K3 carrying Amp33R^{A} which is 33*rinA-ASV*-t-P_{rinA}	This study
pXWAmp30E11	pSB3K3 carrying Amp30E11 which is 30*ECF11*-t-P_{e11}	This study
pXWAmp30E11^{A}	pSB3K3 carrying Amp30E11^{A} which is 30*ECF11*-ASV-t-P_{e11}	This study
pXWAmp31E11	pSB3K3 carrying Amp31E11 which is 31*ECF11*-t-P_{e11}	This study
pXWAmp31E11^{A}	pSB3K3 carrying Amp31E11^{A} which is 31*ECF11*-ASV-t-P_{e11}	This study
pXWAmp32E11	pSB3K3 carrying Amp32E11 which is 32*ECF11*-t-P_{e11}	This study
pXWAmp32E11^{A}	pSB3K3 carrying Amp32E11^{A} which is 32*ECF11*-ASV-t-P_{e11}	This study
pXWAmp33E11	pSB3K3 carrying Amp33E11 which is 33*ECF11*-t-P_{e11}	This study
pXWAmp33E11^{A}	pSB3K3 carrying Amp33E11^{A} which is 33*ECF11*-ASV-t-P_{e11}	This study
pXWHgAmp30^{E}	pSB4A3 carrying J109-32*merR*-t-P_{merT}-Amp30^{E}-30*gfp*-t	This study
pXWHgAmp30R	pSB4A3 carrying J109-32*merR*-t-P_{merT}-Amp30R-30*gfp*-t	This study

pXWHgAmp30R^{A}	pSB4A3 carrying J109-32*merR*-t-P$_{merT}$-Amp30R^{A}-30*gfp*-t	This study
pXWHgAmp31R	pSB4A3 carrying J109-32*merR*-t-P$_{merT}$-Amp31R-30*gfp*-t	This study
pXWHgAmp31R^{A}	pSB4A3 carrying J109-32*merR*-t-P$_{merT}$-Amp31R^{A}-30*gfp*-t	This study
pXWHgAmp32R	pSB4A3 carrying J109-32*merR*-t-P$_{merT}$-Amp32R-30*gfp*-t	This study
pXWHgAmp32R^{A}	pSB4A3 carrying J109-32*merR*-t-P$_{merT}$-Amp32R^{A}-30*gfp*-t	This study
pXWHgAmp33R	pSB4A3 carrying J109-32*merR*-t-P$_{merT}$-Amp33R-30*gfp*-t	This study
pXWHgAmp33R^{A}	pSB4A3 carrying J109-32*merR*-t-P$_{merT}$-Amp33R^{A}-30*gfp*-t	This study
pXWHgAmp30E11	pSB4A3 carrying J109-32*merR*-t-P$_{merT}$-Amp30E11-30*gfp*-t	This study
pXWHgAmp30E11^{A}	pSB4A3 carrying J109-32*merR*-t-P$_{merT}$-Amp30E11^{A}-30*gfp*-t	This study
pXWHgAmp31E11	pSB4A3 carrying J109-32*merR*-t-P$_{merT}$-Amp31E11-30*gfp*-t	This study
pXWHgAmp31E11^{A}	pSB4A3 carrying J109-32*merR*-t-P$_{merT}$-Amp31E11^{A}-30*gfp*-t	This study
pXWHgAmp32E11	pSB4A3 carrying J109-32*merR*-t-P$_{merT}$-Amp32E11-30*gfp*-t	This study
pXWHgAmp32E11^{A}	pSB4A3 carrying J109-32*merR*-t-P$_{merT}$-Amp32E11^{A}-30*gfp*-t	This study
pXWHgAmp33E11	pSB4A3 carrying J109-32*merR*-t-P$_{merT}$-Amp33E11-30*gfp*-t	This study
pXWHgAmp33E11^{A}	pSB4A3 carrying J109-32*merR*-t-P$_{merT}$-Amp33E11^{A}-30*gfp*-t	This study
pXW115HgAmp30E11	pSB3K3 carrying J115-32*merR*-t-P$_{merT}$-30*ECF11*-t-P$_{e11}$-34*mCherry*-t	This study
pXW115HgAmp31E11	pSB3K3 carrying J115-32*merR*-t-P$_{merT}$-31*ECF11*-t-P$_{e11}$-34*mCherry*-t	This study
pXW115HgAmp32E11	pSB3K3 carrying J115-32*merR*-t-P$_{merT}$-32*ECF11*-t-P$_{e11}$-34*mCherry*-t	This study
pXW115HgAmp33E11	pSB3K3 carrying J115-32*merR*-t-P$_{merT}$-33*ECF11*-t-P$_{e11}$-34*mCherry*-t	This study
pXW115HgAmp33E16	pSB3K3 carrying J115-32*merR*-t-P$_{merT}$-33*ECF16*-t-P$_{e16}$-34*mCherry*-t	This study
pXW115HgAmp30E17	pSB3K3 carrying J115-32*merR*-t-P$_{merT}$-30*ECF17*-t-P$_{e17}$-34*mCherry*-t	This study
pXW115HgAmp32E17	pSB3K3 carrying J115-32*merR*-t-P$_{merT}$-32*ECF17*-t-P$_{e17}$-34*mCherry*-t	This study
pXW115HgAmp33E17	pSB3K3 carrying J115-32*merR*-t-P$_{merT}$-33*ECF17*-t-P$_{e17}$-34*mCherry*-t	This study
pXW115HgAmp33E20	pSB3K3 carrying J115-32*merR*-t-P$_{merT}$-33*ECF20*-t-P$_{e20}$-34*mCherry*-t	This study
pXW115HgAmp30E22	pSB3K3 carrying J115-32*merR*-t-P$_{merT}$-30*ECF22*-t-P$_{e22}$-34*mCherry*-t	This study
pXW115HgAmp32E22	pSB3K3 carrying J115-32*merR*-t-P$_{merT}$-32*ECF22*-t-P$_{e22}$-34*mCherry*-t	This study
pXW115HgAmp33E22	pSB3K3 carrying J115-32*merR*-t-P$_{merT}$-33*ECF22*-t-P$_{e22}$-34*mCherry*-t	This study
pXW117As(RS-E11)	pSB4A3 carrying J117-30*arsR*-t-P$_{arsR}$-Amp30^{E}-Amp31E11^{A}-30*gfp*-t	This study
pXW117As(RS-E11-RinA)	pSB4A3 carrying J117-30*arsR*-t-P$_{arsR}$-Amp30^{E}-Amp31E11^{A}-Amp33R-30*gfp*-t	This study
pXW117As(RS-E11-RinA)2	pSB4A3 carrying J117-30*arsR*-t-P$_{arsR}$-Amp30^{E}-Amp31E11^{A}-33*rinA*-t	This study
pXW117As(RS-RinA)	pSB4A3 carrying J117-30*arsR*-t-P$_{arsR}$-Amp30^{E}-Amp30R^{A}-30*gfp*-t	This study
pXW117As(RS-RinA-E11)	pSB4A3 carrying J117-30*arsR*-t-P$_{arsR}$-Amp30^{E}-Amp30R^{A}-Amp33E11-30*gfp*-t	This study
pXW117As(RS-RinA-E11)2	pSB4A3 carrying J117-30*arsR*-t-P$_{arsR}$-Amp30^{E}- Amp30R^{A}-33*E11*-t	This study

pXW109Hg(RS-RinA)	pSB4A3 carrying J109-32*merR*-t-P_{merT}-Amp30^{E}-Amp30R-30*gfp*-t	This study
pXW109Hg(RS-RinA-E11)2	pSB4A3 carrying J109-32*merR*-t-P_{merT}-Amp30^{E}-Amp30R^{A}-33*E11*-t	This study
pXW109Hg(RS-E11)	pSB4A3 carrying J109-32*merR*-t-P_{merT}-Amp30^{E}-Amp31E11^{A}-30*gfp*-t	This study
pXW109Hg(RS-E11-RinA)2	pSB4A3 carrying J109-32*merR*-t-P_{merT}-Amp30^{E}-Amp31E11^{A}-30*rinA-ASV*-t	This study
pXWPrinA-gfp2	pSB1K3 carrying P_{rinA}-30*gfp*-t	This study
pXWP11-gfp2	pSB1K3 carrying P_{e11}-30*gfp*-t	This study
pXW109ABS53(RS)	pXWAmp30^{E}-gfp encoded by J109-30*arsR*-t-P_{arsR}-ABS53	This study
pXW109ABS58(RS)	pXWAmp30^{E}-gfp encoded by J109-30*arsR*-t-P_{arsR}-ABS58	This study
pXW109ABS62(RS)	pXWAmp30^{E}-gfp encoded by J109-30*arsR*-t-P_{arsR}-ABS62	This study
pXW109ABS67(RS)	pXWAmp30^{E}-gfp encoded by J109-30*arsR*-t-P_{arsR}-ABS67	This study
pXW109ABS84(RS)	pXWAmp30^{E}-gfp encoded by J109-30*arsR*-t-P_{arsR}-ABS84	This study
pXW109As(RS-RinA)	pXWJ109ParsR encoding Amp30^{E}-Amp30R-30*gfp*-t	This study
pXWAmp(RS-RinA)-gfp	pSB3K3 carrying Amp30^{E}-Amp30R-30*gfp*-t	This study
pXW109ABS53(RS-RinA)	pXWAmp(RS-RinA)-gfp encoded by J109-30*arsR*-t-P_{arsR}-ABS53	This study
pXW109ABS58(RS-RinA)	pXWAmp(RS-RinA)-gfp encoded by J109-30*arsR*-t-P_{arsR}-ABS58	This study
pXW109ABS62(RS-RinA)	pXWAmp(RS-RinA)-gfp encoded by J109-30*arsR*-t-P_{arsR}-ABS62	This study
pXW109ABS67(RS-RinA)	pXWAmp(RS-RinA)-gfp encoded by J109-30*arsR*-t-P_{arsR}-ABS67	This study
pXW109ABS84(RS-RinA)	pXWAmp(RS-RinA)-gfp encoded by J109-30*arsR*-t-P_{arsR}-ABS84	This study
BBa_K1319004	pSB1C3 carrying *TEV*	*
pXWJ109As(RS-te-AAV)	pXWJ109As carrying Amp30^{E}-30*gfp-te-AAV*-t	This study
pXWJ109As(RS-AAV)	pXWJ109As carrying Amp30^{E}-30*gfp-AAV*-t	This study
PXWABS62(E11-TEV)	pSB4A3 carrying P_{arsR}-ABS62-Amp32E11-32*malE-te-TEV*-t	This study
As0	pXW109ABS67(RS-RinA) in pSB4A3	This study
As1	pXWJ117AsAmp30^{E}2	This study
As2	pXW109ABS84(RS) in pSB4A3	This study
As3	pBW411J117-arsR	Wang et al. 2015
As4	pSB4A3 carrying J115-30*arsR*-t-P_{arsR}-30*gfp*-t	This study
As5	pBW414J105-arsR	Wang et al. 2015
As6	pSB3K3 carrying J109-30*arsR*-t-P_{arsR}-ABS84-*30gfp*-t	This study
As7	pSB3K3 carrying J117-30*arsR*-t-P_{arsR}-ABS84-*30gfp*-t	This study

*: http://parts.igem.org/Part:BBa_K1319004

Appendix Table 2.3: List of oligonucleotides used in this study.

Primer (Set)	Sequence (5'– 3')	Usage
scar_r30_arsR_fwd_P	TACTAGAGATTAAAGAGGAGAAATACTAGATGTCATTTCTGTTACCCATC	For replacing J101 to J117 in the arsenic sensor
J117_E-X_rev	GCTAGCACAATCCCTAGGACTGAGCTAGCTGTCAACTCTAGAAGCGGCCGCGAATTC	
J109-r30-arsR_fwd	CGCTTCTAGAGTTTACAGCTAGCTCAGTCCTAGGGACTGTGCTAGCTACTAGAGATTAAAGAGGAGAAATACTAGATGTCATTTCTGTTACCCATCC	For replacing J101 to J109 in the arsenic sensor
VR	ATTACCGCCTTTGAGTGAGC	
scar_r32_merR_fwd_P	TACTAGAGTCACACAGGAAAGTACTAGATGGAAAATAATTTGGAAAACCTGACCATTG	For replacing J101 to J114/J109 in the mercury sensor
J114_E-X_rev	GCTAGCATTGTACCTAGGACTGAGCTAGCCATAAACTCTAGAAGCGGCCGCGAATTC	
J109_E-X_rev	GCTAGCACAGTCCCTAGGACTGAGCTAGCTGTAAACTCTAGAAGCGGCCGCGAATTC	
Phrpl_opt_fwd_P	CGGTTATTGCTATAGGGCTTGTACTACTAGAGATTAAAGAGGAGAAATACTAG ATGCGTAAAGGAGAAGAAC	For optimising Amp30^{C}
Phrpl_opt_rev	TGCCAGCTTGAAAATAGAAATAACTGATTGATTTATAAAG	
hrpR_B0030	CGTCTAGAGATTAAAGAGGAGAAATACTAGATGAGTACAGGCATCGATAAG	For amplifying 30*hrpR*-30*hrpS*-t-P$_{hrpL}^{E}$
S-P_mPhrpl5_rev	GGGCTGCAGCGGCCGCTACTAGTAGTACAAGCCCTATAGCAATAAC	
E-X_r30_rinA_p80a_fwd	CCGGAATTCGCGGCCGCTTCTAGAGATTAAAGAGGAGAAATACTAGATGACTAAAAAGAAATATGGATTAAAATTATC	For amplifying *rinA_p80α*, and adding various RBS, paired with primer VR
E-X_r31_rinA_p80a_fwd	CCGGAATTCGCGGCCGCTTCTAGAGTCACACAGGAAACCTACTAGATGACTAAAAAGAAATATGGATTAAAATTATC	
E-X_r32_rinA_p80a_fwd	CCGGAATTCGCGGCCGCTTCTAGAGTCACACAGGAAAGTACTAGATGACTAAAAAGAAATATGGATTAAAATTATC	
E-X_r33_rinA_p80a_fwd	CCGGAATTCGCGGCCGCTTCTAGAGTCACACAGGACTACTAGATGACTAAAAAGAAATATGGATTAAAATTATC	
VF2	TGCCACCTGACGTCTAAGAA	For adding an *ASV* tag to *rinA_p80α*
SpeI_ASV_RinA_p80α_rev	CGCTACTAGTATTATTAAACTGATGCAGCGTAGTTTTCGTCGTTTGCAGCAGGCCTCTCTTCCCCTAACTCTTCCG	

E-X_r30_ECF11_987_fwd	CCGGAATTCGCGGCCGCTTCTAGAGATTAAAGAGGAGAAATACTAGATGATGAGCGATAGTCCGCAG	For amplifying *ECF11_987*, adding RBS30 and removing the *ASV* tag
SpeI_TAA_ECF11_987_rev	GCTACTAGTATTATTACAGCTGTTCGGTATGCATGC	
E-X_r31_ECF11_987_fwd	CCGGAATTCGCGGCCGCTTCTAGAGTCACACAGGAAACCTACTAGATGATGAGCGATAGTCCGCAG	For changing the RBS of *ECF11_987*, paired with primer VR
E-X_r32_ECF11_987_fwd	CCGGAATTCGCGGCCGCTTCTAGAGTCACACAGGAAAGTACTAGATGATGAGCGATAGTCCGCAG	
E-X_r33_ECF11_987_fwd	CCGGAATTCGCGGCCGCTTCTAGAGTCACACAGGACTACTAGATGATGAGCGATAGTCCGCAG	
E-X_P11_3726_fwd	AATTCGCGGCCGCTTCTAGAGGCCTCCACACCGCTCGTCACATCCTGTGATCCACTCTTCATCCCGCTAC GTAACACCTCTTA	For constructing P_{ecf11_3726} by oligonucleotides annealing
SpeI_P11_3726_rev	CTAGTAAGAGGTGTTACGTAGCGGGATGAAGAGTGGATCACAGGATGTGACGAGCGGTGTGGAGGCCTCT AGAAGCGGCCGCG	
ABS_S-P_fwd	GTCATATATGTTTTTGACTTATCCGCTTCTACTAGTAGCGGCCGCTGCAG	For inserting extra ABS downstream of promoter P_{arsR}
partial ParsR_rev_53	TTATTGCAGGTAGTGTCTCTCTTC	
partial ParsR_rev_58	TTATAAGTTTGCAGGTAGTGTCTCTCTTC	
partial ParsR_rev_62	TTATGTGTAAGTTTGCAGGTAGTGTCTCTCTTC	
partial ParsR_rev_67	TTAACGAATGTGTAAGTTTGCAGGTAGTGTCTCTCTTC	
Partial ParsR_rev_69	TTAACGAATGTGTAAGTGCTTGCAGGTAGTGTCTCTCTTC	
Partial ParsR_rev_71	TTAACGAATGTGTAAGTGCAGTTGCAGGTAGTGTCTCTCTTC	
Partial ParsR_rev_73	TTAACGAATGTGTAAGTGCAGAGTTGCAGGTAGTGTCTCTCTTC	
Partial ParsR_rev_75	TTAACGAATGTGTAAGTGCAGAGGATTGCAGGTAGTGTCTCTCTTC	
Partial ParsR_rev_77	TTAACGAATGTGTAAGTGCAGAGGAAGTTGCAGGTAGTGTCTCTCTTC	
Partial ParsR_rev_79	TTAACGAATGTGTAAGTGCAGAGGAAGGTTTGCAGGTAGTGTCTCTCTTC	
Partial ParsR_rev_81	TTAACGAATGTGTAAGTGCAGAGGAAGGTAATTGCAGGTAGTGTCTCTCTTC	
Partial ParsR_rev_83	TTAACGAATGTGTAAGTGCAGAGGAAGGTAATATTGCAGGTAGTGTCTCTCTTC	
Partial ParsR_rev_84	TTAACGAATGTGTAAGTGCAGAGGAAGGTAATAGTTGCAGGTAGTGTCTCTCTTC	
EX_r32malE_fwd	TTTCTGGAATTCGCGGCCGCTTCTAGAGTCACACAGGAAAGTACTAGATGATCGAAGAAGGTAAACTGGT AATC	For amplifying *malE* from *E. coli* MG1655 genome
TEVlink_malE_rev	GCCACCCTGAAAATACAGGTTTTCGGATCCCTTGGTGATACGAGTCTGCG	

TEVlink_TEV_fwd	ATCCGAAAACCTGTATTTTCAGGGTGGCGGCGAAAGCCTGTTCAAAGG	For amplifying *TEV*
PS_TAA_TEV_rev	GCCGGACTGCAGCGGCCGCTACTAGTATTATTAATTCATCAGTTGGGTTGCTTC	
AAV-gfp_rev	CGCTACTAGTATTATTAAACTGCGGCAGCGTAGTTTTCGTCGTTTGCAGCAGGCCTTTTGTATAGTTCATCCATGCC	For adding an *AAV* tag to *gfp*, paired with primer VF2
TEVlink_ASV_fwd	TATTTTCAGGGTGGCGCTGCAAACGACGAAAACTAC	For inserting TEV protease cleavage site between *gfp* and *AAV* tag
TEVlink_gfp_rev	CAGGTTTTCGGATCCTTTGTATAGTTCATCCATGCCATG	
linker_TEVsite_fwd	GGCACCGGCGGCTCCGGCGGCTCCGAAAACCTGTATTTTCAGG	For adding extra linker before the TEV protease cleavage site
linker_gfp_rev	GCCGGAGCCGCCGGAGCCGCCTTTGTATAGTTCATCCATG	
linker_malE_rev	GCCGGAGCCGCCGGAGCCGCCCTTGGTGATACGAGTCTG	
2ParsR35_rev1	GAAGCGGATATCACAAAAACATATATGAC	For changing σ^{70} binding sites in P_{arsR}
2ParsR35_rev2	GAAGCGGATAACACAAAAACATATATGAC	
2ParsR35_rev3	GAAGCGGATAGGTCAAAAACATATATGAC	
2ParsR35_rev4	GAAGCGGATACGTCAAAAACATATATGAC	
2ParsR35_rev5	GAAGCGGATACATCAAAAACATATATGAC	
2ParsR10_fwd1	GAAGAGATATGTTACCTGCAATACTAGAGAAAG	
2ParsR10_fwd2	GAAGAGAGATACTACCTGCAATACTAGAGAAAG	
2ParsR10_fwd3	GAAGAGAGATAATACCTGCAATACTAGAGAAAG	
2ParsR10_fwd4	GAAGAGATATAGTACCTGCAATACTAGAGAAAG	
2ParsR10_fwd5	GAAGAGATATAATACCTGCAATACTAGAGAAAG	
X_35cABS-RBS_fwd	GCTTCTAGAGTTGTGAGCTAGCTCTAAGTCATATATGTTTTTGACTTAACTAGAGAAAGAGGAGAAATAC	For engineering arsenic inducible promoter via the second design, paired with VR
X_35dABS-RBS_fwd	GCTTCTAGAGTTTACAGCTAGCTCTAAGTCATATATGTTTTTGACTTAACTAGAGAAAGAGGAGAAATAC	
X_35eABS-RBS_fwd	GCTTCTAGAGTAGACAGCTAGCTCTAAGTCATATATGTTTTTGACTTAACTAGAGAAAGAGGAGAAATAC	
X_103_35ABS10-RBS_fwd	GCTTCTAGAGCTGATAAGTCATATATGTTTTTGACTTATGCTAGCTACTAGAGAAAGAGGAGAAATAC	
X_105_35ABS10-RBS_fwd	GCTTCTAGAGCTGATAAGTCATATATGTTTTTGACTTATGCTAGCTACTAGAGAAAGAGGAGAAATAC	
EX_rbs30_spacer_fwd	CTGGAATTCGCGGCCGCTTCTAGAGATTAAAGAGGAGAAATACTAG	
EX_rbs32_spacer_fwd	CTGGAATTCGCGGCCGCTTCTAGAGTCACACAGGAAAGTACTAG	

EX_rbs33_spacer_fwd	CTGGAATTCGCGGCCGCTTCTAGAGTCACACAGGACTACTAG	To amplify ECFs, but only bind to the RBS and spacer sequence, paired with VR
RemovePstIE16_rev2	CTTGCAGCAGTAATGCTTGCAGCTGGCTTTCACGGGTGCTC	To remove the PstI sites in *ECF16* and *ECF*20
RemovePstIE16_fwd2	CAAGCATTACTGCTGCAAGGTCTGGCAGGCGATACCTTTGCATATCGTC	
RemovePstIE20_rev	GATAATATTGCAGAACAATTGCTTCACGCTG	
RemovePstIE20_fwd	CAATTGTTCTGCAATATTATCAAGAACTGAGC	
ECF16_3622_fwd	CGTCTAGAGATTAAAGAGGAGAAATACTAGATGCAGCGTACCAATAGCCAGGATGTTCTG	For constructing P_{ecf16_3622}, P_{ecf17_up1691}, P_{ecf20_992} and P_{ecf22_up1147} by oligonucleotides annealing
ECF16_3622_rev	GCCTGCAGCGGCCGCTACTAGTATTAACGATCTTCATCATGACCTTTACCAC	
ECF17_1691_fwd	CGTCTAGAGATTAAAGAGGAGAAATACTAGATGGCACGTGTTAGCGGTG	
ECF17_1691_rev	GCCTGCAGCGGCCGCTACTAGTATTAACGGGTAACACCCAGTTCTTG	
ECF20_992_fwd	CGTCTAGAGATTAAAGAGGAGAAATACTAGATGAATGAAACCGATCCTGATCTGG	
ECF20_992_rev	GCCTGCAGCGGCCGCTACTAGTATTACGGTTTGCGACGACCG	
ECF22_4450_fwd	CGTCTAGAGATTAAAGAGGAGAAATACTAGATGCCGCAGCAGACCGATC	
ECF22_4450_rev	GCCTGCAGCGGCCGCTACTAGTATACAGTTCTGCACGAATACGTGC	
EX_PmerT_fwd	AATTCGCGGCCGCTTCTAGAGTTCCATATCGCTTGACTACGTACATGAGTACGGAAGTAAGGTTACGCTATCCAATCCTA	For constructing P_{merT} by oligonucleotides annealing
SpeI_PmerT_rev	CTAGTAGGATTGGATAGCGTAACCTTACTTCCGTACTCATGTACGTAGTCAAGCGATATGGAACTCTAGAAGCGGCCGCG	
E_X_r30_lacZ_fwd	CCGGAATTCGCGGCCGCTTCTAGAGATTAAAGAGGAGAAATACTAGATGACCATGATTACGGATTCACTGGCC	For amplifying *lacZ* from *E. coli* MG1655 genome
S_lacZ_rev	GCTACTAGTATTATTTTTGACACCAGACCAACTGGTAATGGTAGCGACCGGCGCTCAGCTGAAACTCCGCCGATACTGACGGGCTCC	
hrpR_fwd_seq	GTGGATCTGCGGATCATTGC	Sequencing primers binding to *hrpR* and *hrpS*
hrpR_Rev_seq	CTTCAATTCGCGGATGTTGC	
hrpS_fwd_seq	CGACGATCAAGCTGGATATC	
hrpS_Rev_seq	CAAACATCGGGAACGGGAAC	
lacZ_fwd_seq	ATTATTTGCCCGATGTACGCGC	

lacZ_rev_seq	TGTAAACGGGGATACTGACG	Sequencing primers binding to *lacZ*
lacZ-N_rev_seq	TTAAGTTGGGTAACGCCAGG	
GFPmut3b_fwd_seq	CAAGAGTGCCATGCCCGAAG	Sequencing primers binding to *gfp*
GFPmut3b_rev_seq	AATGGTTGTCTGGTAAAAGG	
mC.N_seq.R	CCCATGGTCTTCTTCTGCATTACG	Sequencing primer binding to *mCherry*
malE_fwd_seq	TCAACGGCCCGTGGGCATGG	Sequencing primer binding to *malE*
PmerT88_fwd	GCCGTACATGAGTACGGAAGTAAG	Sequencing primer binding to P_{merT}

Appendix Table 2.4: Best fits for the characterised responses of the various sensors including circuit amplified ones in this study.

(Data shown are with 95% confidence bounds)**.**

Sensor	Target	*k* (a.u.)	α	*n*	K_M (μM)	R^2	Used in
J101-P_{arsR}-*gfp*	$NaAsO_2$	4450 ± 232	1.49E-16 ± 0.0058	1.2616 ± 0.0793	1.4547 ± 0.1347	0.9950	
J117-P_{arsR}-*gfp*	$NaAsO_2$	11816 ± 444	1.36E-16 ± 0.0056	1.1573 ± 0.0611	0.9922 ± 0.0741	0.9958	Fig. 3.2b
J109-P_{arsR}-*gfp*	$NaAsO_2$	14450 ± 433	1.28E-16 ± 0.0092	0.9849 ± 0.5333	0.3711 ± 0.0250	0.9951	
J109-P_{arsR}-*gfp*	$NaAsO_2$	6784 ± 861	0.0355 ± 0.0060	1.8232 ± 0.1285	0.1171 ± 0.0159	0.9960	
J109-P_{arsR}-Am30^{C}-*gfp*	$NaAsO_2$	73437 ± 857	0.0136 ± 0.0046	3.0965 ± 0.1505	0.0311 ± 0.0005	0.9976	Fig. 3.4a
J109-P_{arsR}-Am30^{E}-*gfp*	$NaAsO_2$	90182 ± 1276	0.0216 ± 0.0071	3.2062 ± 0.1924	0.0249 ± 0.0006	0.9949	
1-layer (RS) (3K3)	$NaAsO_2$	57780 ± 641	1.34E-16 ± 0.0052	3.6256 ± 0.2155	0.1313 ± 0.0020	0.9968	
1-layer (RS) (4A3)	$NaAsO_2$	27032 ± 213	9.84E-14 ± 0.0044	2.7021 ± 0.0892	0.0739 ± 0.0011	0.9982	
2-layer (RS-E11) (4A3)	$NaAsO_2$	61119 ± 1073	0.0141 ± 0.0099	2.4483 ± 0.1734	0.0651 ± 0.0021	0.9837	Fig. 3.6b and c
3-layer (RS-E11-RinA) (4A3)	$NaAsO_2$	24747 ± 1181	0.0249 ± 0.0288	4.4035 ± 1.7038	0.0630 ± 0.0039	0.9345	
3-layer (RS-E11-RinA) (4A3) + P_{rinA}-*gfp* (1K3)	$NaAsO_2$	159615 ± 2047	0.0201 ± 0.0078	2.9568 ± 0.1780	0.0566 ± 0.0013	0.9927	
1-layer (RS) (3K3)	$NaAsO_2$	52765 ± 496	1.57E-16 ± 0.0045	2.6934 ± 0.1011	0.1048 ± 0.0017	0.9984	
2-layer (RS-RinA) (4A3)	$NaAsO_2$	59252 ± 714	1.27E-16 ± 0.0069	2.0815 ± 0.0915	0.0516 ± 0.0012	0.9975	
3-layer (RS-RinA-E11) (4A3)	$NaAsO_2$	29060 ± 1213	0.0744 ± 0.0329	4.4556 ± 1.0859	0.0254 ± 0.0018	0.9595	Fig. 3.7a
3-layer (RS-RinA-E11) (4A3) + P_{e11}-*gfp* (1K3)	$NaAsO_2$	86155 ± 564	0.0185 ± 0.0044	3.1806 ± 0.1033	0.0360 ± 0.0004	0.9991	
J115-P_{merT}-*gfp*	$HgCl_2$	125550 ± 84184	0.0019 ± 0.0021	2.1518 ± 0.3067	5.9008 ± 2.9466	0.9960	
J114-P_{merT}-*gfp*	$HgCl_2$	498702 ± 1135836	0.0011 ± 0.0030	1.7335 ± 0.3343	11.6056 ± 19.9495	0.9909	Fig. 3.8a
J109-P_{merT}-*gfp*	$HgCl_2$	114742 ± 24442	0.0140 ± 0.0145	1.1537 ± 0.2248	1.9089 ± 0.7700	0.9620	

0-layer (3K3)	$HgCl_2$	5812566 ± 1.54E+09	2.29E-05 ± 0.0061	0.7371 ± 0.1881	3332 ± 1206132	0.9767	Fig. 3.8b
1-layer (RS) (3K3)	$HgCl_2$	66167 ± 4281	0.0003 ± 0.0080	0.7944 ± 0.0689	0.0785 ± 0.0150	0.9937	
2-layer (RS-RinA) (4A3)	$HgCl_2$	28742 ± 1450	1.15E-16 ± 0.0256	0.8437 ± 0.1073	0.0063 ± 0.0010	0.9766	
3-layer (RS-RinA-E11) (4A3) + P_{e11}-*gfp* (1K3)	$HgCl_2$	37953 ± 2273	0.0289 ± 0.0314	1.3688 ± 0.2388	0.0023 ± 0.0003	0.9610	
P_{arsR}	$NaAsO_2$	28453 ± 1272	0.0076 ± 0.0340	2.8000 ± 0.5663	0.0323 ± 0.0025	0.9610	Fig. 3.10b
P_{arsR}-ABS53	$NaAsO_2$	31309 ± 1462	0.0040 ± 0.0347	3.0000 ± 0.6471	0.0372 ± 0.0031	0.9550	
P_{arsR}-ABS62	$NaAsO_2$	45757 ± 3840	0.0007 ± 0.0550	3.7900 ± 1.7851	0.0732 ± 0.0102	0.8594	
P_{arsR}-ABS67	$NaAsO_2$	61165 ± 6425	0.0005 ± 0.0593	4.5000 ± 2.6894	0.1537 ± 0.0265	0.7932	
P_{arsR}-ABS84	$NaAsO_2$	28554 ± 1409	0.0006 ± 0.0217	3.6000 ± 0.7946	0.3036 ± 0.0223	0.9543	
P_{arsR}	$NaAsO_2$	155453 ± 36624	0.2173 ± 0.1772	0.8539 ± 0.3800	0.0325 ± 0.0143	0.8374	Fig. 3.10c
P_{arsR}-ABS62	$NaAsO_2$	148783 ± 8583	0.0548 ± 0.0391	3.0000 ± 0.7696	0.0427 ± 0.0045	0.9316	
P_{arsR}-ABS67	$NaAsO_2$	147403 ± 6807	0.0086 ± 0.0259	3.7107 ± 0.7860	0.0824 ± 0.0064	0.9542	
P_{arsR}-ABS84	$NaAsO_2$	125605 ± 6732	0.0060 ± 0.0249	4.0000 ± 1.2845	0.1348 ± 0.0097	0.9461	
P_{arsR}-Amp30^{E}-*gfp*	$NaAsO_2$	38445 ± 1315	0.0253 ± 0.0011	1.8624 ± 0.0542	0.0455 ± 0.0023	0.9914	Fig. 3.10e
P_{arsR}-Amp30^{E}-*gfp-AAV*	$NaAsO_2$	7112 ± 1523	0.0034 ± 0.0008	7.0156 ± 0.8048	0.1259 ± 0.0107	0.6592	
P_{arsR}-Amp30^{E}-*gfp-te-AAV* + *TEV*	$NaAsO_2$	52492 ± 3257	0.0013 ± 0.0001	2.7916 ± 0.0989	0.0337 ± 0.0020	0.9583	
As0	$NaAsO_2$	37044 ± 606	0.0100 ± 0.0074	2.5378 ± 0.1571	0.1191 ± 0.0033	0.9955	Fig. 4.2a
As1	$NaAsO_2$	23677 ± 143	1.32E-18 ± 0.0031	2.9446 ± 0.0736	0.0955 ± 0.0010	0.9993	
As2	$NaAsO_2$	21680 ± 883	1.41E-18 ± 0.0150	1.3723 ± 0.1337	0.7721 ± 0.0629	0.9872	
As3	$NaAsO_2$	12544 ± 591	2.22E-18 ± 0.0098	1.1179 ± 0.0859	1.3721 ± 0.1334	0.9932	
As4	$NaAsO_2$	2194 ± 99	1.46E-13 ± 0.0182	0.8797 ± 0.0760	1.6260 ± 0.1794	0.9892	
As5	$NaAsO_2$	5244 ± 383	2.22E-18 ± 0.0078	0.9919 ± 0.0797	10.0419 ± 1.5943	0.9938	

As6	$NaAsO_2$	1370 ± 73	0.0100 ± 0.0148	0.9319 ± 0.0825	3.4553 ± 0.4241	0.9902	
As7	$NaAsO_2$	1223 ± 32	2.22E-18 ± 0.0042	1.1774 ± 0.0452	7.3883 ± 0.3791	0.9983	
As0	$NaAsO_2$	42768 ± 196	0.0174 ± 0.0019	2.3960 ± 0.0377	0.1428 ± 0.0011	0.9997	Fig. 4.8a–d
As0	Na_2HAsO_4	53994 ± 2447	0.0101 ± 0.0021	1.9765 ± 0.0852	0.6593 ± 0.0350	0.9986	
As1	$NaAsO_2$	25220 ± 143	0.0002 ± 0.0029	3.1556 ± 0.0792	0.1047 ± 0.0010	0.9993	
As1	Na_2HAsO_4	25624 ± 629	2.13E-16 ± 0.0051	2.3416 ± 0.1296	0.3329 ± 0.0106	0.9963	
As2	$NaAsO_2$	19898 ± 500	1.17E-16 ± 0.0075	1.3943 ± 0.0768	1.0761 ± 0.0515	0.9961	
As2	Na_2HAsO_4	19021 ± 1001	2.22E-16 ± 0.0045	1.4646 ± 0.0910	3.4465 ± 0.2910	0.9961	
As3	$NaAsO_2$	12220 ± 368	1.24E-16 ± 0.0068	1.1985 ± 0.0635	1.3379 ± 0.0794	0.9967	
As3	Na_2HAsO_4	10812 ± 270	4.67E-16 ± 0.0038	1.4398 ± 0.0561	2.2922 ± 0.0981	0.9982	

Appendix Table 2.5: Arsenic analysis of groundwater samples collected from Bangladesh.

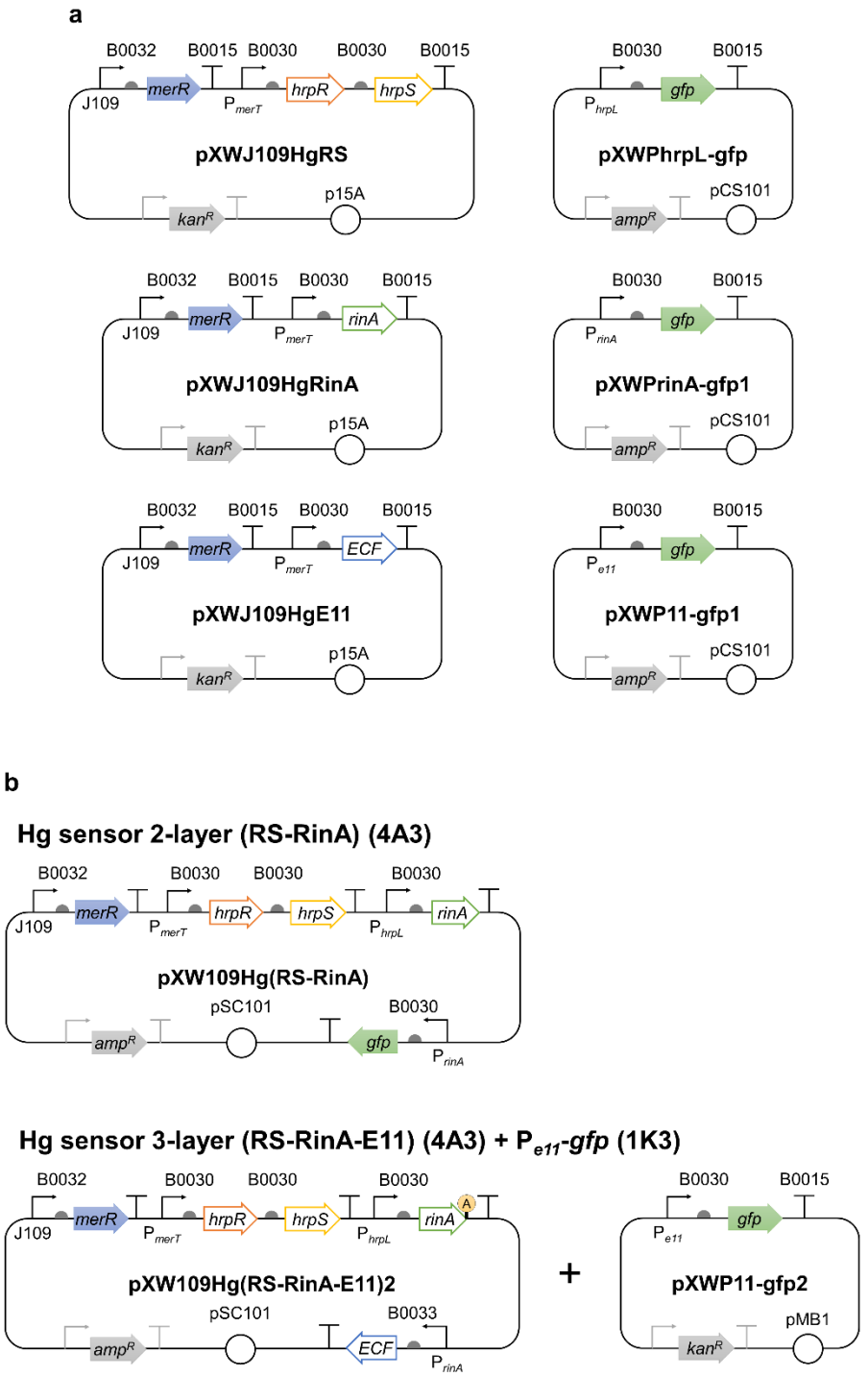
	As / 75 (m.w.)*		Sb / 121 (m.w.)	
LOD	0.753 ppb		0.004 ppb	
Sample:	ppb	% r.s.d	ppb	% r.s.d
1	49.84	7.27	0.0012	>100.00
2	98.96	0.09	-0.0019	32.01
3	40.47	8.28	0.0157	14.58
4	14.28	3.72	-0.0031	10.41
5	17.82	4.79	-0.0046	10.87
6	39.30	7.05	0.0011	>100.00
7	59.31	6.73	-0.0005	43.30
8	1.41	83.55	-0.0021	23.49
9	55.34	3.72	0.0349	4.88
10	32.31	7.27	0.0254	5.71
11	103.70	1.36	0.0087	8.91
12	90.60	2.76	0.0039	21.83
13	113.30	3.20	0.0113	9.16
14	20.18	14.60	0.0066	18.57
15	351.60	0.14	0.0070	7.79
16	119.80	1.11	0.0146	11.19
17	119.90	2.64	-0.0002	>100.00
18	83.30	1.92	0.0092	14.24
19	65.67	1.60	0.0527	6.90
20	10.20	12.38	0.0046	20.54
21	169.00	1.96	0.0216	1.72
22	3.18	14.52	0.0141	2.62

*: values shown are for total arsenic, including both arsenite and arsenate.

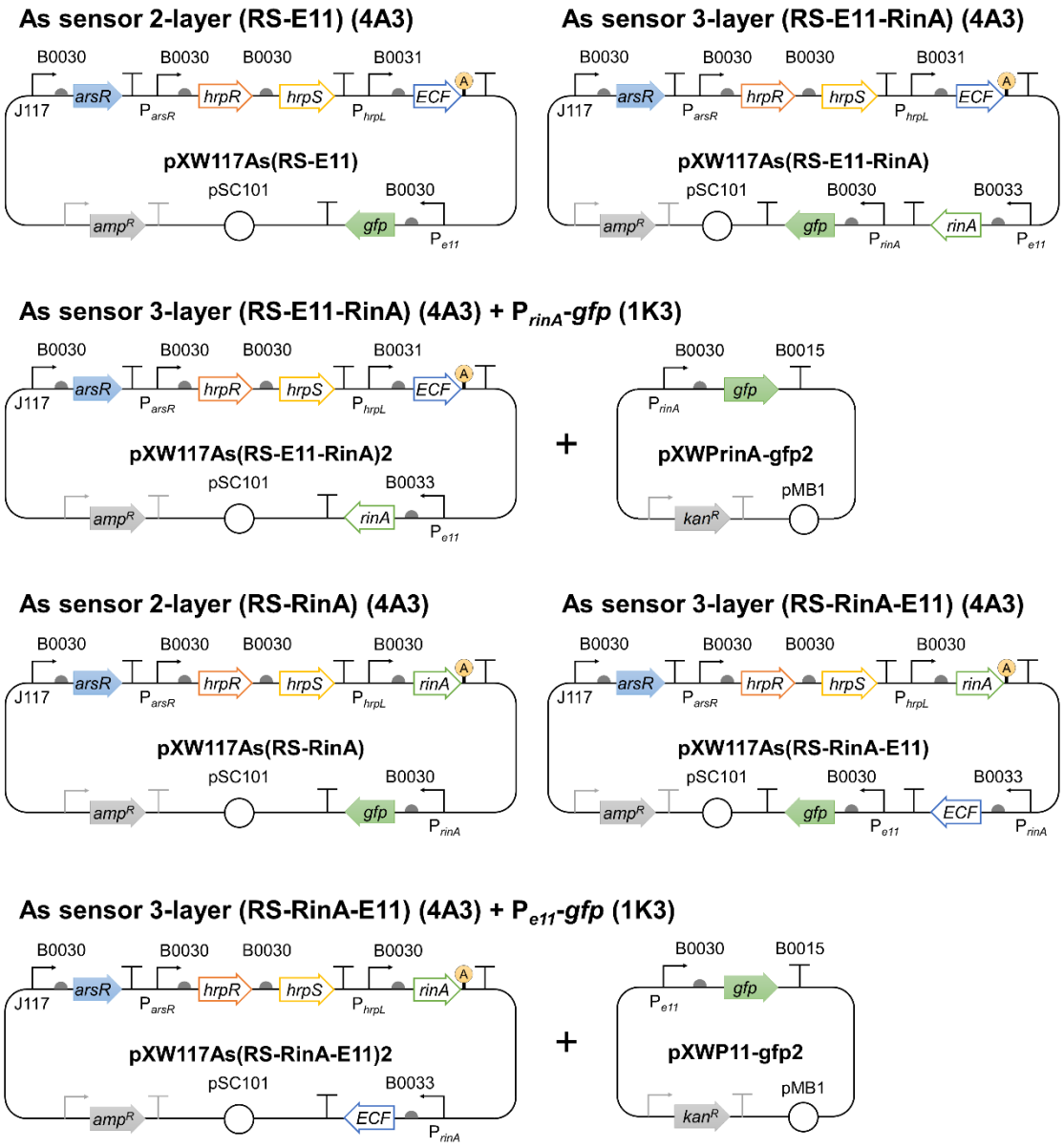
m.w.: molecular weight.

r.s.d: relative standard deviation.

Appendix 2. Figures for Chapter 3

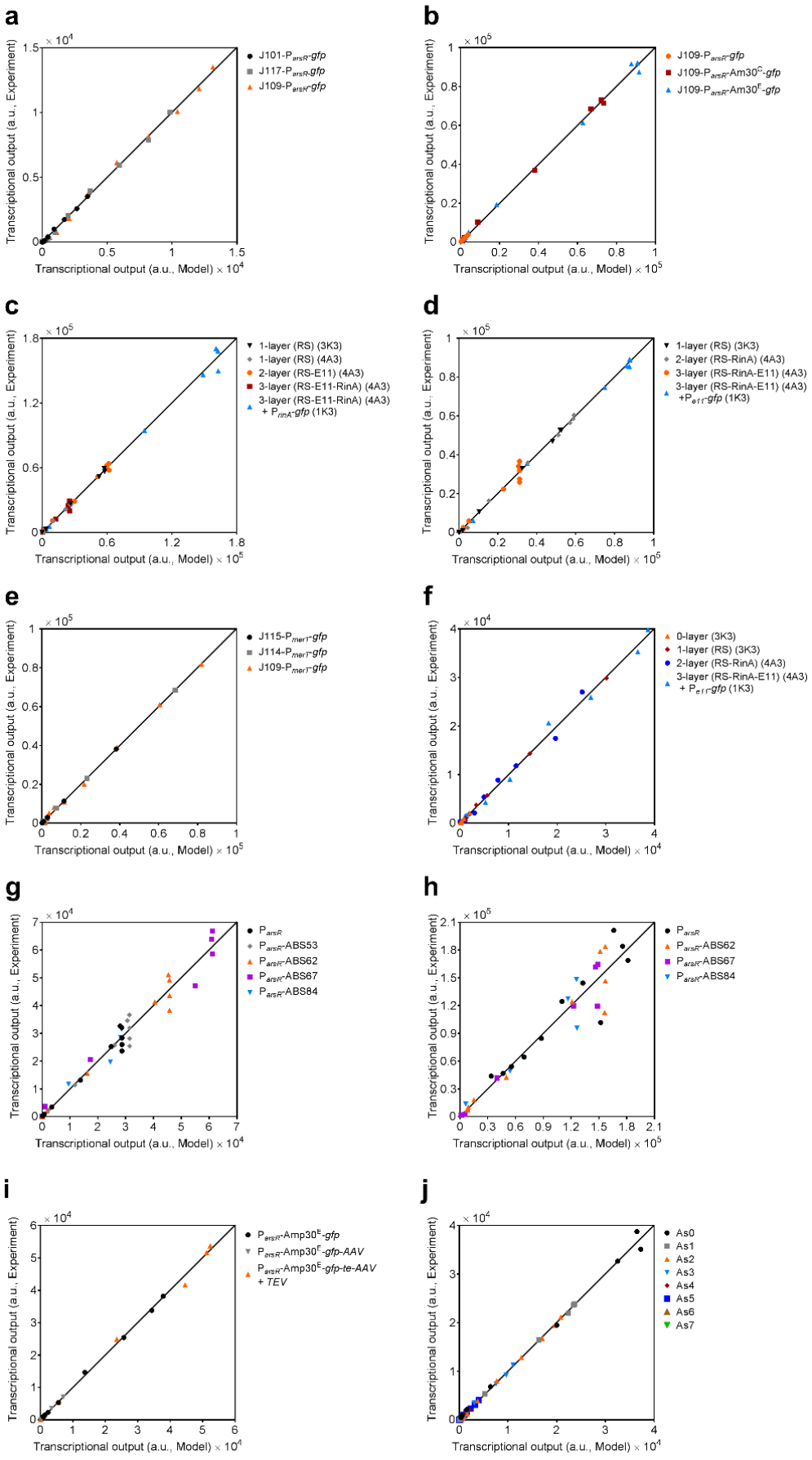


c

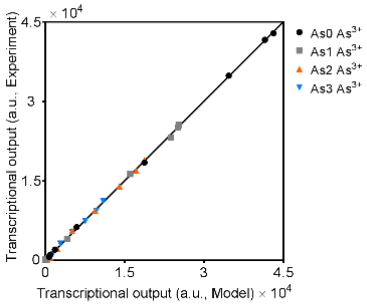


Appendix Figure 3.1: Plasmid maps of sensors with amplifiers.

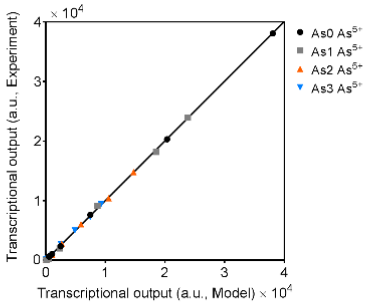
Related to **Figure 3.6a–c**, **Figure 3.7a** and **Figure 3.8c**. Plasmid maps showing the genetic circuits used in the amplifier orthogonality characterisation (**a**), and the multi-layered amplifier tests for mercury (**b**) and arsenic (**c**) sensors. ⊤ represents terminator B0015. P_{hrpL} is the optimised *hrpL* promoter with the consensus σ^{54} motif sequences.


a
Transcriptional output (a.u., Experiment)
Transcriptional output (a.u., Model)
J101-P$_{arsR}$-gfp
J117-P$_{arsR}$-gfp
J109-P$_{arsR}$-gfp
b
J109-P$_{arsR}$-gfp
J109-P$_{arsR}$-Am30^{C}-gfp
J109-P$_{arsR}$-Am30^{F}-gfp
c
1-layer (RS) (3K3)
1-layer (RS) (4A3)
2-layer (RS-E11) (4A3)
3-layer (RS-E11-RinA) (4A3)
3-layer (RS-E11-RinA) (4A3) + P$_{rinA}$-gfp (1K3)
d
1-layer (RS) (3K3)
2-layer (RS-RinA) (4A3)
3-layer (RS-RinA-E11) (4A3)
3-layer (RS-RinA-E11) (4A3) +P$_{e11}$-gfp (1K3)
e
J115-P$_{merT}$-gfp
J114-P$_{merT}$-gfp
J109-P$_{merT}$-gfp
f
0-layer (3K3)
1-layer (RS) (3K3)
2-layer (RS-RinA) (4A3)
3-layer (RS-RinA-E11) (4A3) + P$_{e11}$-gfp (1K3)
g
P$_{arsR}$
P$_{arsR}$-ABS53
P$_{arsR}$-ABS62
P$_{arsR}$-ABS67
P$_{arsR}$-ABS84
h
P$_{arsR}$
P$_{arsR}$-ABS62
P$_{arsR}$-ABS67
P$_{arsR}$-ABS84
i
P$_{arsR}$-Amp30^{E}-gfp
P$_{arsR}$-Amp30^{E}-gfp-AAV
P$_{arsR}$-Amp30^{E}-gfp-te-AAV + TEV
j
As0
As1
As2
As3
As4
As5
As6
As7

k



l



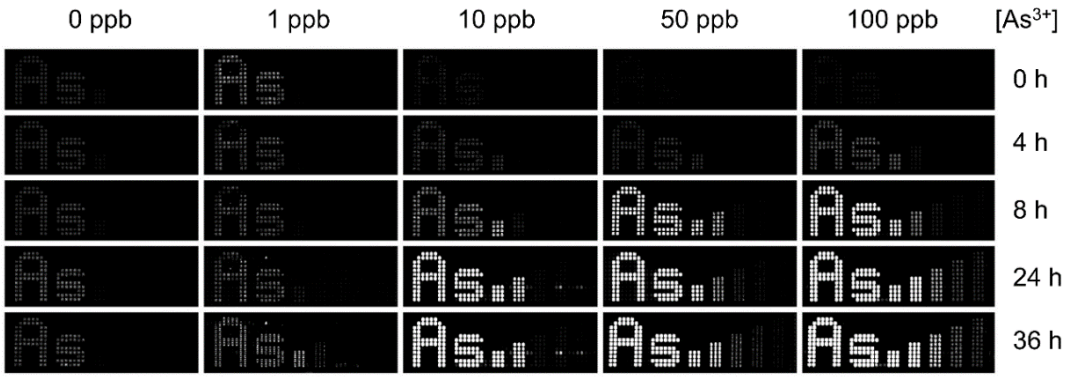
m

	Sensor	Pearson's r		Sensor	Pearson's r
Fig. a	J101-P*arsR*-*gfp*	0.9993	Fig. d	1-layer (*E*) (3K3)	0.9997
	J117-P*arsR*-*gfp*	0.9989		2-layer (*E-R*) (4A3)	0.9994
	J109-P*arsR*-*gfp*	0.9983		3-layer (*E-R-E11*) (4A3)	0.9799
Fig. b	J109-P*arsR*-*gfp*	1.0000		3-layer (*E-R-E11*) (4A3) +P*11*-*gfp* (1K3)	0.9997
	J109-P*arsR*-Am30^{C}-*gfp*	0.9994	Fig. f	0-layer (3K3)	0.9978
	J109-P*arsR*-Am30^{E}-*gfp*	0.9988		1-layer (*E*) (3K3)	0.9999
Fig. c	1-layer (*E*) (3K3)	0.9994		2-layer (*E-R*) (4A3)	0.9922
	1-layer (*E*) (4A3)	0.9998		3-layer (*E-R-E11*) (4A3) + P*11*-*gfp* (1K3)	0.9967
	2-layer (*E-E11*) (4A3)	0.9980	Fig. g	P*arsR*	0.9814
	3-layer (*E-E11-R*) (4A3)	0.9834		P*arsR*-ABS53	0.9796
	3-layer (*E-E11-R*) (4A3) + P*rinA*-*gfp* (1K3)	0.9973		P*arsR*-ABS62	0.9894
Fig. e	J115-P*merT*-*gfp*	1.0000		P*arsR*-ABS67	0.9926
	J114-P*merT*-*gfp*	0.9999		P*arsR*-ABS84	0.9865
	J109-P*merT*-*gfp*	0.9989	Fig. j	As0	0.9979
Fig. h	P*arsR*	0.9300		As1	0.9998
	P*arsR*-ABS62	0.9660		As2	0.9996
	P*arsR*-ABS67	0.9860		As3	0.9978
	P*arsR*-ABS84	0.9744		As4	0.9971
Fig. i	P*arsR*-Amp30^{E}-*gfp*	0.9997		As5	0.9979
	P*arsR*-Amp30^{E}-*gfp-AAV*	1.0000		As6	0.9989
	P*arsR*-Amp30^{E}-*gfp-linte-AAV* + *TEV*	0.9987		As7	0.9997
Fig. k	As0 As^{3+}	0.9999	Fig. l	As0 As^{5+}	1.0000
	As1 As^{3+}	0.9998		As1 As^{5+}	0.9996
	As2 As^{3+}	0.9995		As2 As^{5+}	0.9998
	As3 As^{3+}	0.9986		As3 As^{5+}	0.9992

Appendix Figure 3.2: Comparing experimental characterisation data with model fitting outcome for the dose-response of sensors induced with arsenic or mercury.

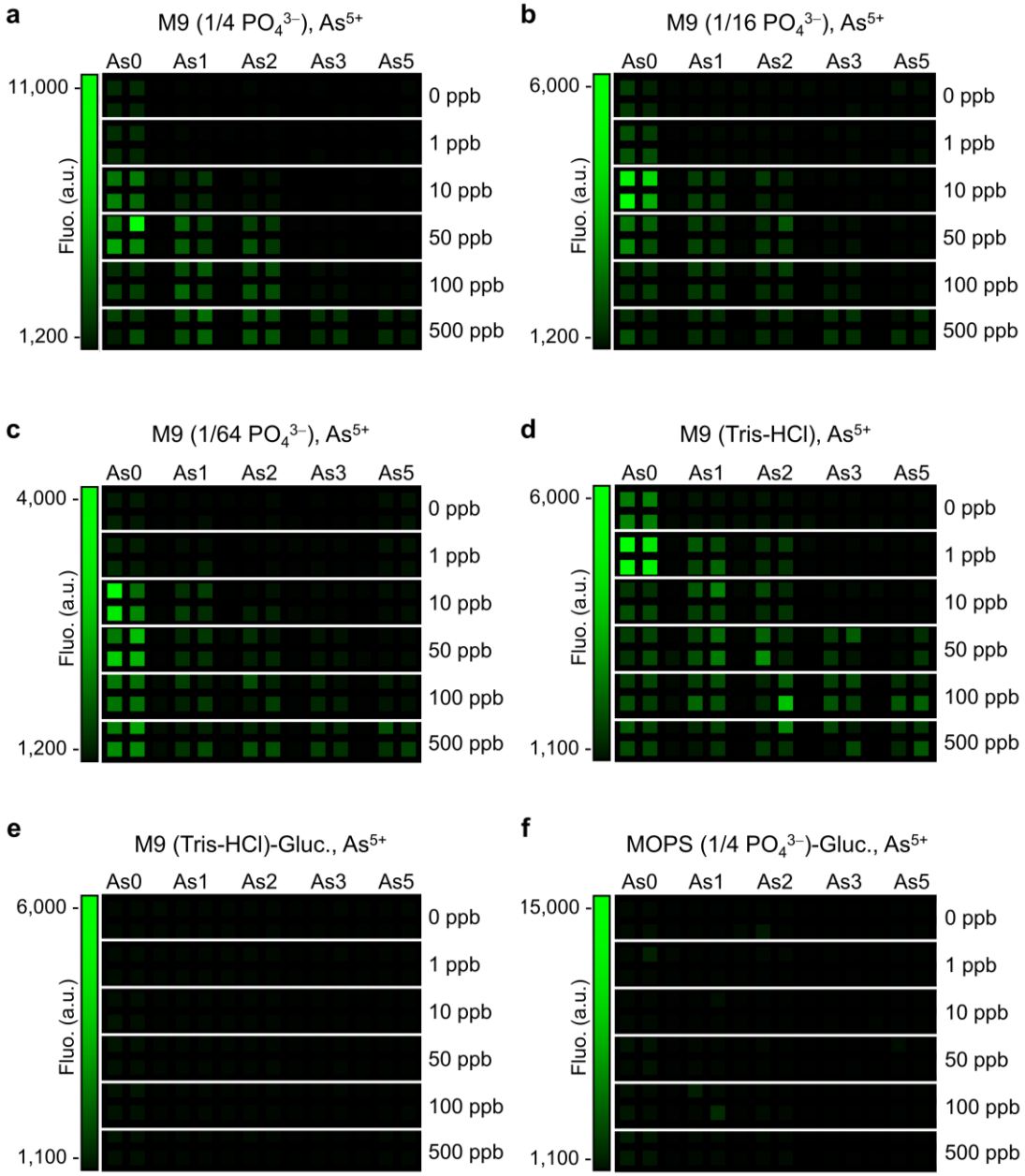
a is related to **Figure 3.2b**, **b** is related to **Figure 3.4a**, **c** is related to **Figure 3.6b**–**c**, **d** is related to **Figure 3.7a**, **e** is related to **Figure 3.8a**, **f** is related to **Figure 3.8c**, **g** is related to **Figure 3.10b**, **h** is related to **Figure 3.10c**, **i** is related to **Figure 3.10d**, **j** is related to **Figure 4.2a**, **k** and **l** are related to **Figure 4.8a**–**d**. Table **m** shows the Pearson correlation coefficient (Pearson's *r*) between the model predicted and experimentally characterised responses of each sensor shown in **a**–**l**.

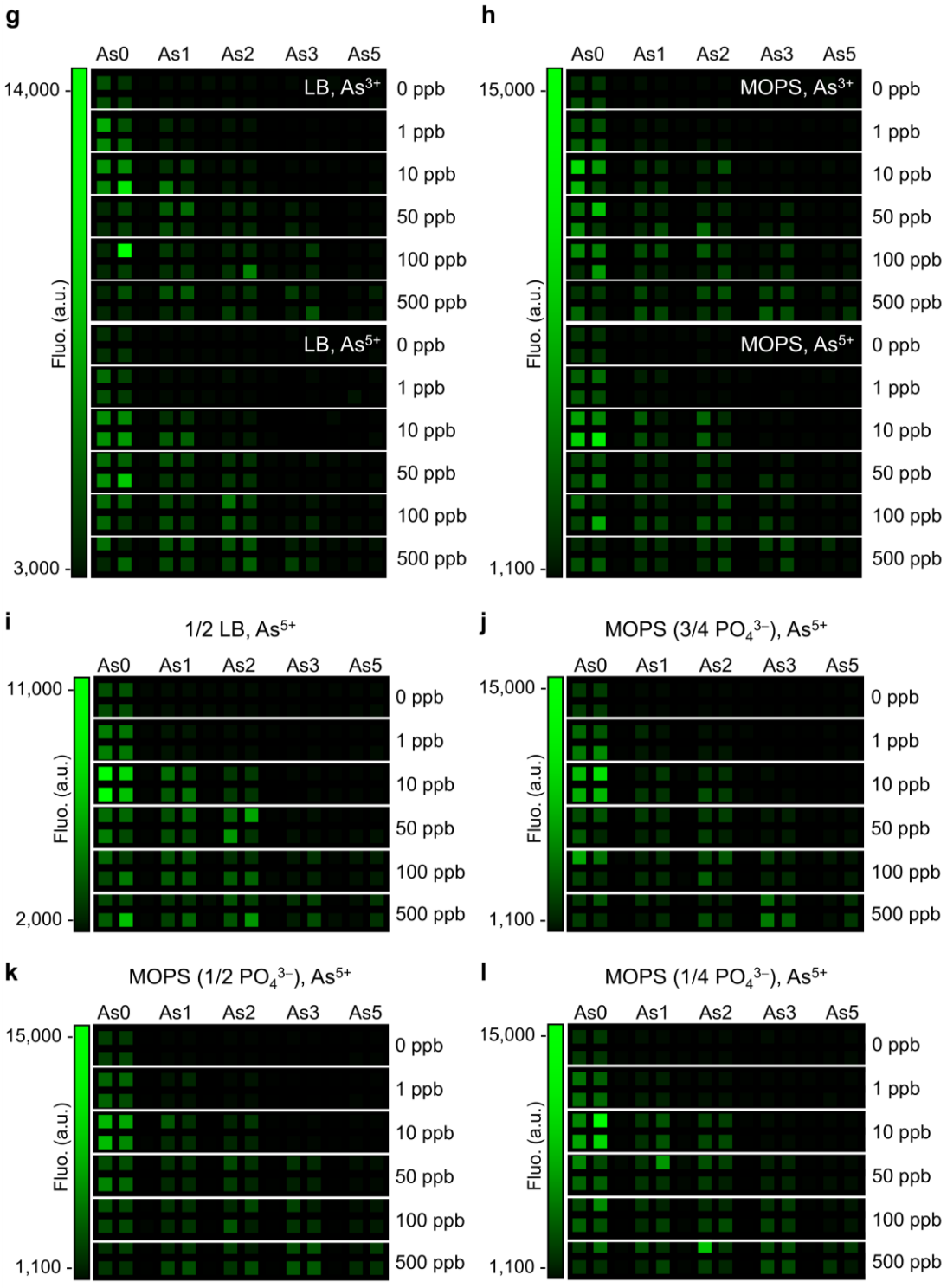
Appendix 3. Figures for Chapter 4



Appendix Figure 4.1: Microbial sensor array display enabled by microfluidic encapsulation.

Related to **Figure 4.3c–e**. Nikon fluorescent microscopy images of microfluidic encapsulation-enabled sensor array under various arsenic induction levels over time.


a
M9 (1/4 PO4 3−), As5+
As0 As1 As2 As3 As5
11,000
1,200
Fluo. (a.u.)
0 ppb
1 ppb
10 ppb
50 ppb
100 ppb
500 ppb
b
M9 (1/16 PO4 3−), As5+
As0 As1 As2 As3 As5
6,000
1,200
Fluo. (a.u.)
0 ppb
1 ppb
10 ppb
50 ppb
100 ppb
500 ppb
c
M9 (1/64 PO4 3−), As5+
As0 As1 As2 As3 As5
4,000
1,200
Fluo. (a.u.)
0 ppb
1 ppb
10 ppb
50 ppb
100 ppb
500 ppb
d
M9 (Tris-HCl), As5+
As0 As1 As2 As3 As5
6,000
1,100
Fluo. (a.u.)
0 ppb
1 ppb
10 ppb
50 ppb
100 ppb
500 ppb
e
M9 (Tris-HCl)-Gluc., As5+
As0 As1 As2 As3 As5
6,000
1,100
Fluo. (a.u.)
0 ppb
1 ppb
10 ppb
50 ppb
100 ppb
500 ppb
f
MOPS (1/4 PO4 3−)-Gluc., As5+
As0 As1 As2 As3 As5
15,000
1,100
Fluo. (a.u.)
0 ppb
1 ppb
10 ppb
50 ppb
100 ppb
500 ppb

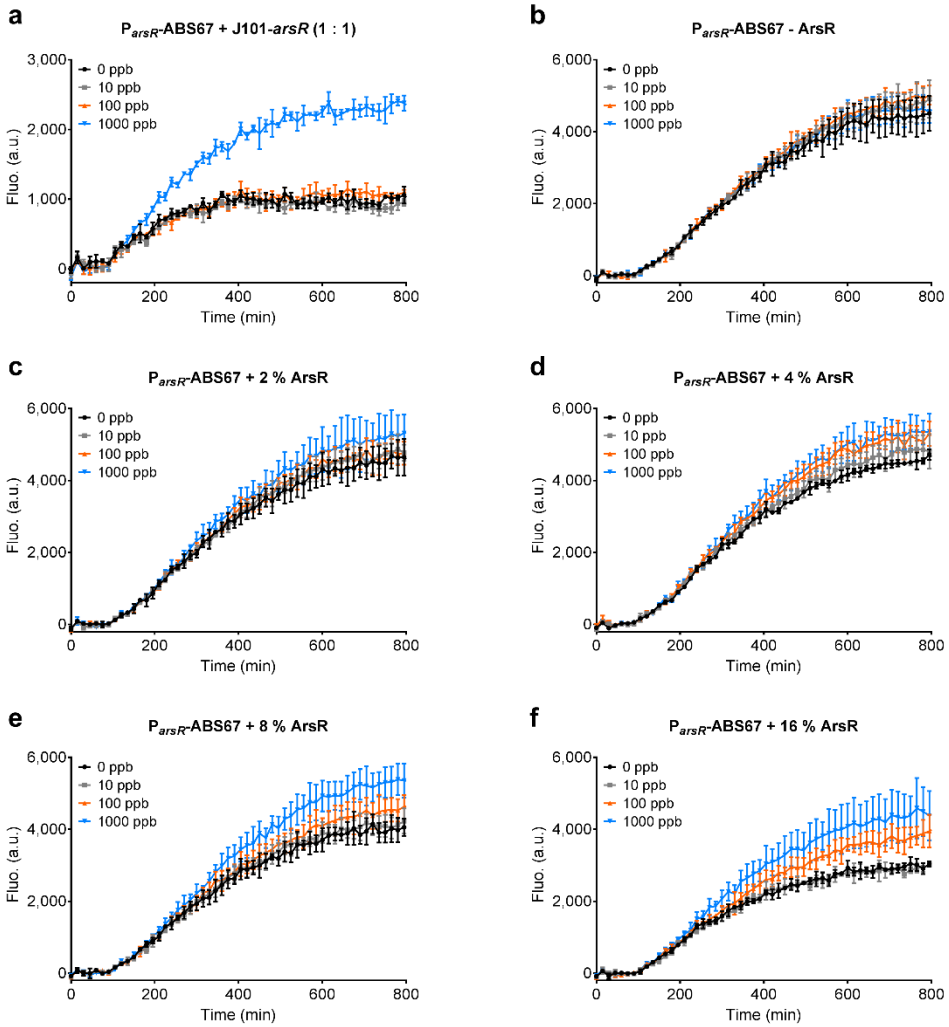


Appendix Figure 4.2: Response comparison of agarose gel-entrapped arsenic sensors in the presence of arsenite or arsenate.

(**a–d**) Improving arsenic sensors' responses to arsenate (Na_2HAsO_4) by reducing the phosphate levels in M9 medium to 1/4 (**a**), 1/16 (**b**) and 1/64 (**c**) of the original phosphate concentration, or by substituting the phosphate-based buffer by a Tris-HCl-based buffer (**d**, see Section 2.5.2). Sensor incubated with

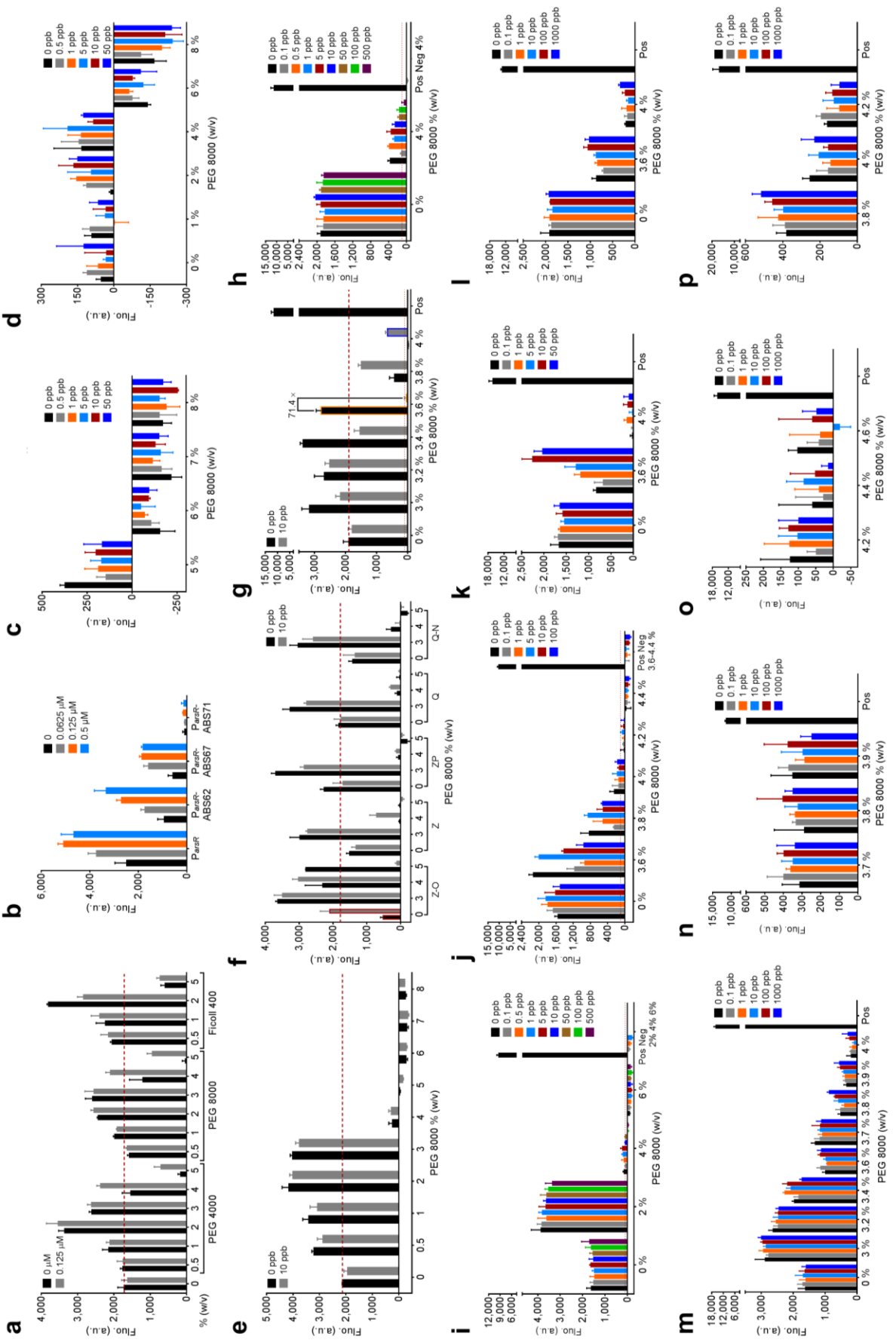
M9 (Tris-HCl) had good sensitivities to arsenate. However, the overall fluorescence was too weak to generate sufficiently contrasted images by a cell phone. More importantly, As1 and As2 were not properly distinguished, as was As3 and As5. (**e,f**) Characterisation of arsenic sensors As0–3 and As5 in response to arsenate (Na_2HAsO_4) in M9 (**e**) or MOPS rich medium (**f**) with glucose (Gluc.) as the carbon source. Agarose gel entrapped sensors did not work under this condition. We observed that the entrapped cells did not grow in the media with glucose as the carbon source. (**g,h**) Characterisation of arsenic sensors As0–3 and As5 in response to the same concentration of arsenite ($NaAsO_2$) or arsenate (Na_2HAsO_4) in LB (**g**) or MOPS rich medium (**h**). (**i**) Characterisation of arsenic sensors As0–3 and As5 in response to arsenate (Na_2HAsO_4) in 2-fold diluted LB, which was used to reduce background fluorescence. (**j–l**) Improving arsenic sensors' response to arsenate (Na_2HAsO_4) by reducing the phosphate levels in MOPS medium to 3/4 (**j**), 1/2 (**k**) and 1/4 (**l**) of the original concentration. MOPS medium with 1/4 PO_4^{3-} gave the best responses for As1 and As5 towards arsenate and the best cell phone images, so it was selected for further tests of water samples. a.u., arbitrary units.

Appendix 4. Figures and Tables for Chapter 5



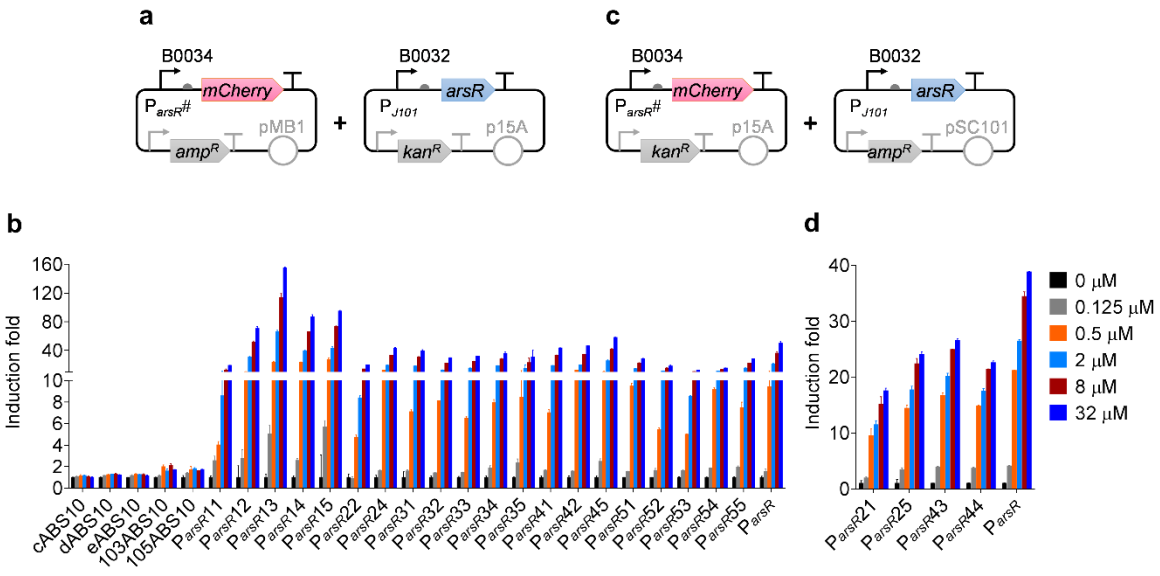
Appendix Figure 5.1: Characterisation of P$_{arsR}$-ABS67 in the presence of ArsR expression.

Related to **Figure 5.3d**. Time-course characterisation of each sensor shown in **Figure 5.3d**. Error bars, s.d. (n = 3, technical replicates). a.u., arbitrary units.


a
b
c
d
e
f
g
h
i
j
k
l
m
n
o
p
Fluo. (a.u.)
PEG 8000 % (w/v)
PEG 4000
PEG 8000
Ficoll 400
PansR
PansR-ABS62
PansR-ABS67
PansR-ABS71
Z-O
Z
ZP
Q
Q-N
Pos
Pos Neg 4%
Pos Neg 3.6-4.4 %
Pos Neg 2% 4% 6%

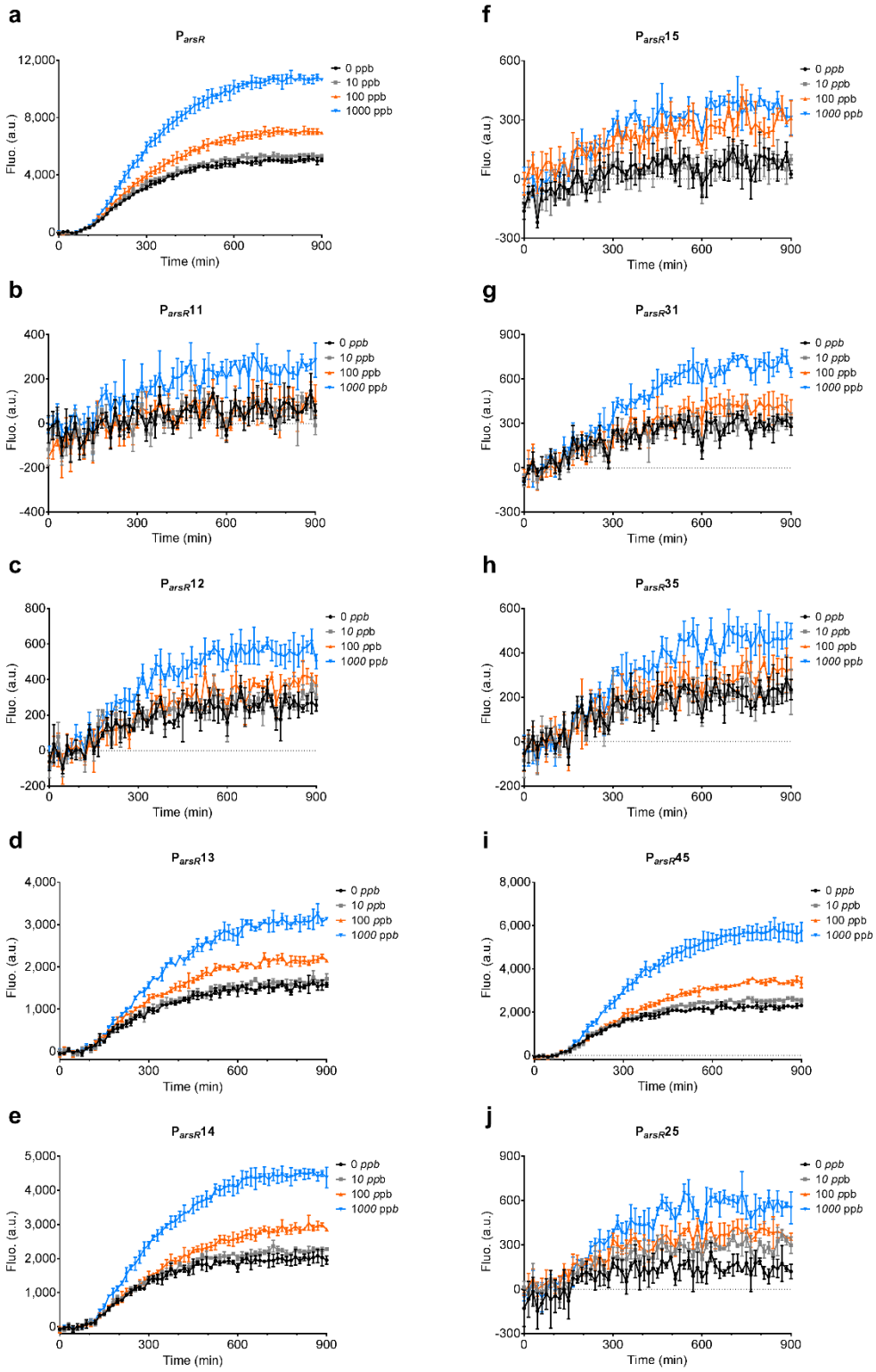
Appendix Figure 5.2: Characterisation of an arsenic sensor with crowding agents in CFS.

Related to **Figure 5.4**. Characterisation of arsenic sensors with crowding agents in CFS. All experiments are for P_{arsR}-ABS67 sensor unless mentioned otherwise in graphs. In **f**, Z represents DNA purified from Zymo kit, Q represents DNA purified from Qiagen kit, and ZP represents DNA purified from Zymo kit but followed by an ethanol precipitation, O represents the PEG 8000 used in **a** – **d**, N represents a new batch of PEG 8000, the others without O or N used PEG 8000 same as in **e**. The final volume of cell-free reaction mixture was 5 µl. All the data were collected after 4 h incubation. Error bars, s.d. ($n \geq 3$, technical replicates). a.u., arbitrary units.



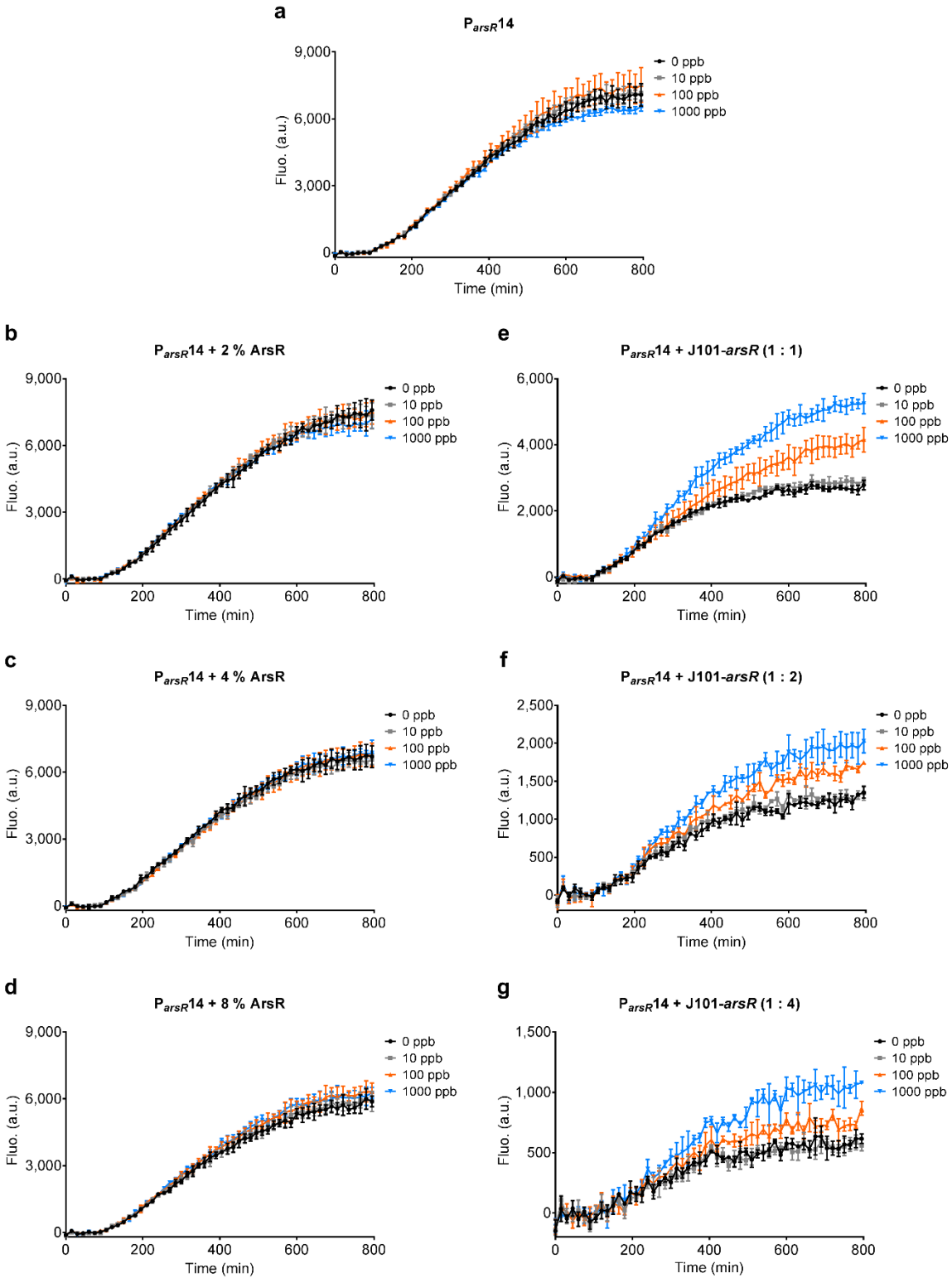
Appendix Figure 5.3: *In vivo* characterisation of engineered arsenic inducible promoters.

Related to **Figure 5.6c,d**. *In vivo* characterisation of engineered arsenic inducible promoters induced with various $NaAsO_2$ concentrations. **a** shows genetic circuits used for testing engineered arsenic inducible promoters in **b**. **c** shows genetic circuits used for testing engineered arsenic inducible promoters in **d**. All the data were collected after 5 h incubation. Error bars, s.d. ($n \geq 2$). a.u., arbitrary units.



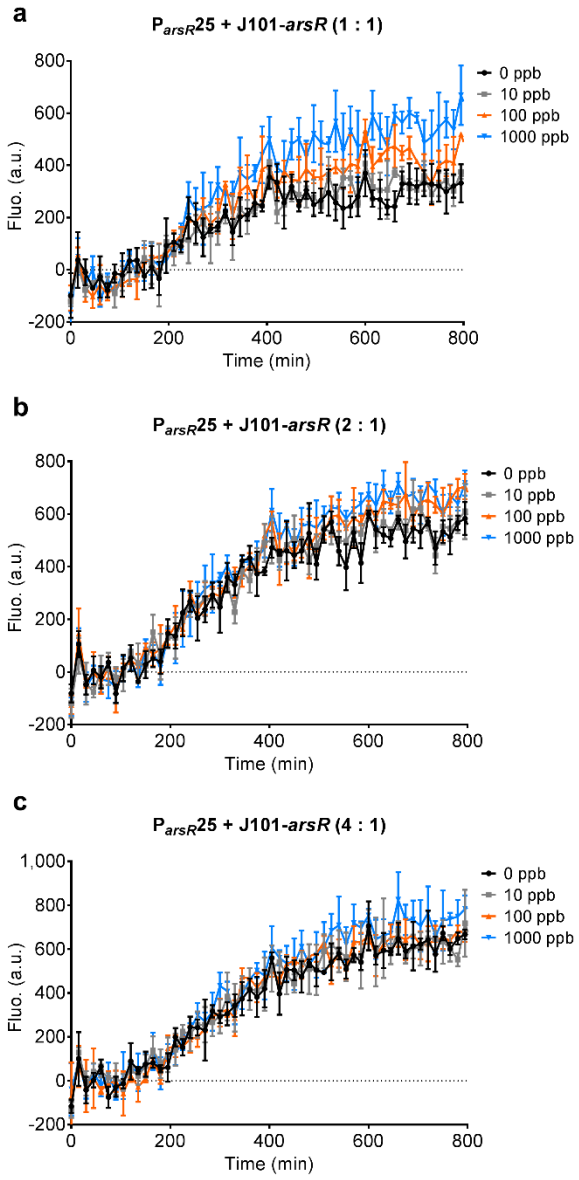
Appendix Figure 5.4: Characterisation of engineered arsenic inducible promoters in CFS.

Related to **Figure 5.7b**. Time-course characterisation of each sensor shown in **Figure 5.7b**. Error bars, s.d. (n = 3, technical replicates). a.u., arbitrary units.



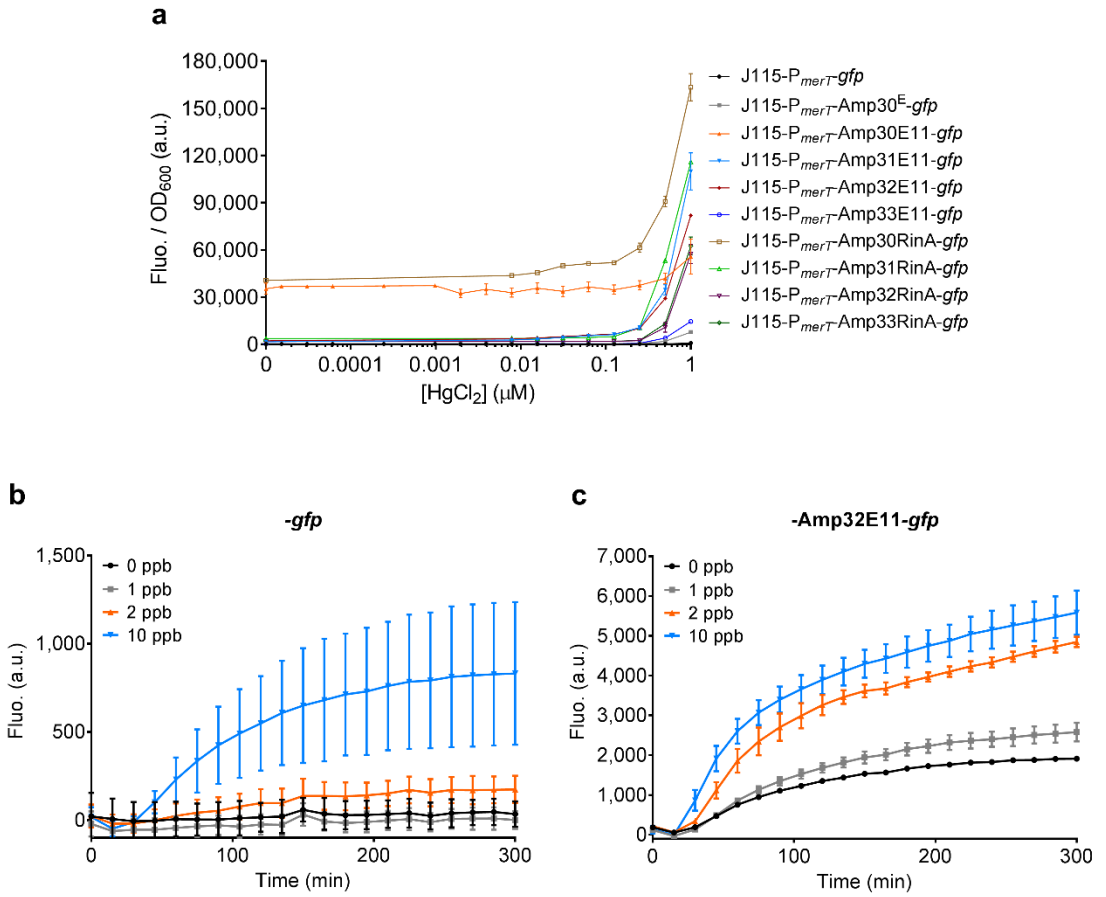
Appendix Figure 5.5: Tuning the background leakage and detection limit of P$_{arsR}$14 in CFS.

Related to **Figure 5.8a**. Time-course characterisation of P$_{arsR}$14 sensor with different levels of ArsR expression. Error bars, s.d. (n = 3, technical replicates). a.u., arbitrary units.



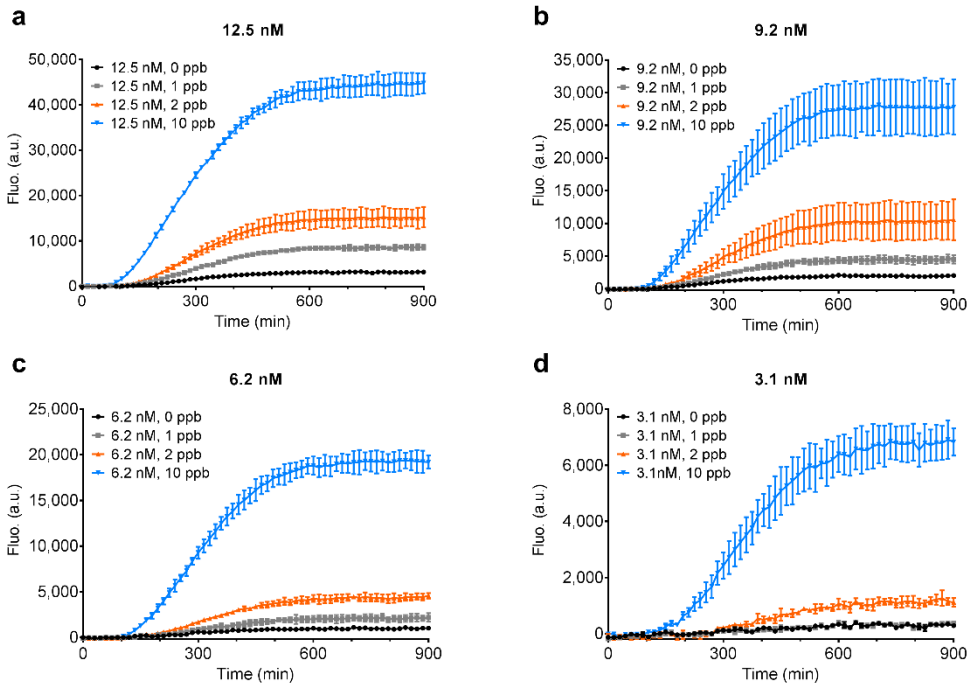
Appendix Figure 5.6: Tuning the background leakage and detection limit of P$_{arsR}$25 in CFS.

Related to **Figure 5.8b**. Time-course characterisation of P$_{arsR}$25 sensor with different levels of ArsR expression. Error bars, s.d. (n = 3, technical replicates). a.u., arbitrary units.



Appendix Figure 5.7: Improving sensor sensitivity and output dynamics by employing amplifiers in CFS.

Related to **Figure 5.10**. (**a**) *In vivo* characterisation of mercury sensors coupled with amplifiers. Data were collected after 6 h induction and incubation. Error bars, s.d. (n ≥ 1). (**b** and **c**) Time-course characterisation of mercury sensors shown in **Figure 5.10c** and **f**. Error bars, s.d. (n = 3, technical replicates). a.u., arbitrary units.



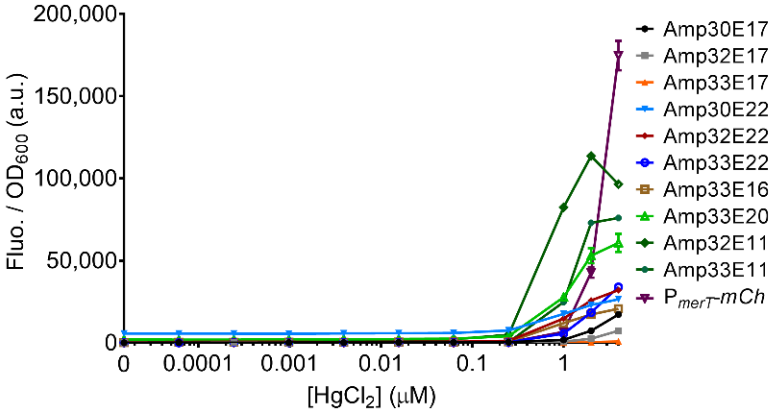
Appendix Figure 5.8: Reducing mercury sensor background leakage by tuning DNA concentration in CFS.

Related to **Figure 5.11b**. Time-course characterisation of the mercury sensor characterised in **Figure 5.11b**. Error bars, s.d. (n = 3, technical replicates). a.u., arbitrary units.

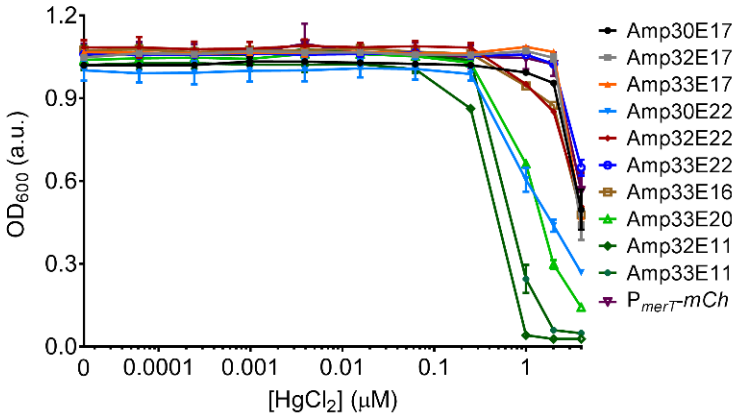
a

	Fluo./OD_{600} (a.u.)					
[Arab.] (mM)	**0**	**0.05**	**0.21**	**0.83**	**3.3**	**13.3**
Amp32E11	10478	46264	39175	32203	26722	26152
Amp33E11	1420	19934	36218	34706	30671	25077
Amp33E16	1883	16750	45375	58549	52976	47943
Amp30E17	-10	1502	8437	10559	9749	9546
Amp32E17	78	1302	4493	5890	5189	4581
Amp33E17	-18	740	2488	3203	2809	2522
Amp33E20	3607	40894	94753	108245	93231	84054
Amp30E22	5779	17810	41730	53985	50316	42982
Amp32E22	923	8826	25773	33329	27871	24972
Amp33E22	91	1741	7373	11392	9998	8636

b

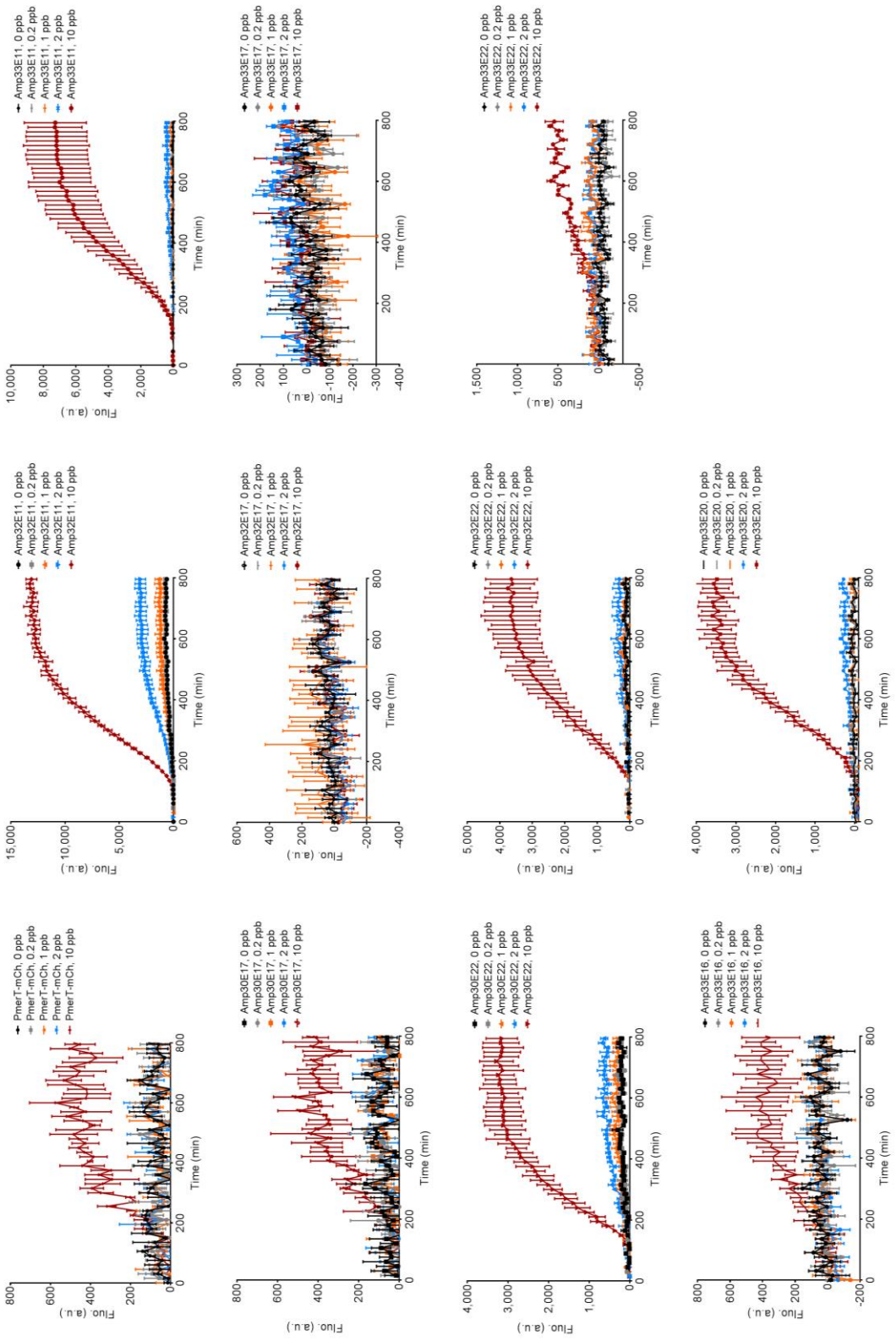


c



Appendix Figure 5.9: *In vivo* characterisation of ECF-based amplifiers.

In vivo characterisation of ECF-based amplifiers using an arabinose (Arab.) inducible promoter (**a**) or using the J115-based mercury sensor module (**b** and **c**). Amp represents amplifier. E11/16/17/20/22 represents ECF11/16/17/20/22. Numbers between Amp and E represent RBS0030/32/33. P_{merT}-*mCh* represents the mercury sensor without amplifier. *mCh*, *mCherry*. All the data were collected after 5 h incubation. Error bars, s.d. ($n \geq 2$). a.u., arbitrary units.



Appendix Figure 5.10: Characterisation of ECF-based amplifiers in CFS.

Related to **Figure 5.12**. Time-course characterisation of the mercury sensors characterised in **Figure 5.12**. Error bars, s.d. (n = 3, technical replicates). a.u., arbitrary units.

Appendix Table 5.1: Summary of major cell-free sensors' LOD and output readout

Sensor	Target	LOD (ppb)*	Output**	Induction fold***	DNA concentration (nM)	Reaction volume (µl)	Receptor expression	Testing condition	Tested in
J101/105/114/117-*arsR*-P_{arsR}-*gfp*	$NaAsO_2$	-	-	-	13.4	10	Yes	30 ℃, overnight	-
J114/117-*arsR* + P_{arsR}-*mCherry*	$NaAsO_2$	-	-	-	8.6 + 9.1	50	Yes	37 ℃, overnight	-
J101-*arsR* + P_{arsR}-*mCherry*	$NaAsO_2$	-	-	-	5.3 + 1.1	50	Yes	37 ℃, overnight	-
J101-*arsR* + P_{arsR}-*mCherry*	$NaAsO_2$	≥ 10	-	-	5.3 + 4.5/2.3/1.1/0.6	50	Yes	37 ℃, overnight	Fig. 5.2
P_{arsR}-ABS67-*gfp*	$NaAsO_2$	10	8,371.00 ± 744.60	1.61 ± 0.22	10.8	10	No	37 ℃, 4 h	Fig. 5.3a
P_{arsR}-ABS67-*gfp*	$NaAsO_2$	10	2,002.00 ± 245.37	1.31 ± 0.10	10.3	5	No	37 ℃, 4 h	Fig. 5.3c
J101-*arsR* + P_{arsR}-ABS67-*gfp*	$NaAsO_2$	100	2090.67 ± 195.41	2.15 ± 0.20	9.2 + 9.2	4	Yes	37 ℃, 8 h	Fig. 5.3d
P_{arsR}-ABS84-*gfp*	$NaAsO_2$	100	244.00 ± 9.89	3.01 ± 0.34	10.8	10	No	37 ℃, 4 h	Fig. 5.3a
J101-*arsR* + P_{arsR}-*mCherry*	$NaAsO_2$	10	4751.33 ± 52.44	1.04 ± 0.01	8.2 + 8.2	4	Yes	37 ℃, 8 h	Fig. 5.7
J101-*arsR* + P_{arsR}11-*mCherry*	$NaAsO_2$	1,000	278.00 ± 85.98	5.12 ± 1.58	8.2 + 8.2	4	Yes	37 ℃, 8 h	
J101-*arsR* + P_{arsR}12-*mCherry*	$NaAsO_2$	1,000	498.67 ± 85.17	2.07 ± 0.35	8.2 + 8.2	4	Yes	37 ℃, 8 h	
J101-*arsR* + P_{arsR}13-*mCherry*	$NaAsO_2$	100	1347.67 ± 30.11	1.37 ± 0.07	8.2 + 8.2	4	Yes	37 ℃, 8 h	

J101-*arsR* + P_{arsR}14-*mCherry*	$NaAsO_2$	100	2070.67 ± 75.16	1.25 ± 0.09	8.2 + 8.2	4	Yes	37 ℃, 8 h	
J101-*arsR* + P_{arsR}15-*mCherry*	$NaAsO_2$	100	152.33 ± 84.33	2.05 ± 0.19	8.2 + 8.2	4	Yes	37 ℃, 8 h	
J101-*arsR* + P_{arsR}25-*mCherry*	$NaAsO_2$	10	287.33 ± 35.01	1.77 ± 0.22	8.2 + 8.2	4	Yes	37 ℃, 8 h	
J101-*arsR* + P_{arsR}31-*mCherry*	$NaAsO_2$	1,000	601.00 ± 30.81	2.11 ± 0.11	8.2 + 8.2	4	Yes	37 ℃, 8 h	
J101-*arsR* + P_{arsR}35-*mCherry*	$NaAsO_2$	1,000	401.33 ± 66.11	1.77 ± 0.29	8.2 + 8.2	4	Yes	37 ℃, 8 h	
J101-*arsR* + P_{arsR}45-*mCherry*	$NaAsO_2$	100	2242.33 ± 97.03	1.32 ± 0.03	8.2 + 8.2	4	Yes	37 ℃, 8 h	
J115-P_{merT}-*gfp*	$HgCl_2$	10	750.00	7.14	6.7	50	Yes	37 ℃, 2 h	
J115-P_{merT}-*rfp*	$HgCl_2$	2	134.50	9.28	6.7	50	Yes	37 ℃, 2 h	Fig. 5.9
J117-P_{merT}-*rfp*	$HgCl_2$	10	124.50	2.24	6.7	50	Yes	37 ℃, 2 h	
J115-P_{merT}-*gfp*	$HgCl_2$	10	548.67 ± 267.50	33.59 ± 16.38	12.5	5	Yes	37 ℃, 2 h	
J115-P_{merT}-Amp30E11-*gfp*	$HgCl_2$	1	2,770.00 ± 61.80	1.14 ± 0.04	12.5	5	Yes	37 ℃, 2 h	
J115-P_{merT}-Amp31E11-*gfp*	$HgCl_2$	1	2,233.67 ± 148.14	1.17 ± 0.02	12.5	5	Yes	37 ℃, 2 h	Fig. 5.10
J115-P_{merT}-Amp32E11-*gfp*	$HgCl_2$	1	1,681.00 ± 118.65	1.24 ± 0.09	12.5	5	Yes	37 ℃, 2 h	
J115-P_{merT}-Amp33E11-*gfp*	$HgCl_2$	10	1279.33 ± 832.86	31.72 ± 20.65	12.5	5	Yes	37 ℃, 2 h	

J115-P*merT*-Amp32E11-*mCherry*	$HgCl_2$	1	2,408.67 ± 120.09	2.41 ± 0.12	12.5	5	Yes	37 ℃, 2 h	
J115-P*merT*-Amp32E11-*mCherry*	$HgCl_2$	1	1,352.67 ± 125.80	1.78 ± 0.16	9.2	5	Yes	37 ℃, 2 h	Fig. 5.11
J115-P*merT*-Amp32E11-*mCherry*	$HgCl_2$	1	452.00 ± 89.06	1.63 ± 0.32	6.2	5	Yes	37 ℃, 2 h	
J115-P*merT*-Amp32E11-*mCherry*	$HgCl_2$	10	1194.00 ± 264.91	30.10 ± 6.68	3.1	5	Yes	37 ℃, 2 h	
J115-P*merT*-*mCherry*	$HgCl_2$	10	499.33 ± 70.74	8.37 ± 1.19	6.2	4	Yes	37 ℃, 12 h	
J115-P*merT*-Amp32E11-*mCherry*	$HgCl_2$	1	1,415.00 ± 338.91	2.00 ± 0.48	6.2	4	Yes	37 ℃, 12 h	
J115-P*merT*-Amp33E11-*mCherry*	$HgCl_2$	10	7,208.67 ± 2,027.94	69.09 ± 19.44	6.2	4	Yes	37 ℃, 12 h	
J115-P*merT*-Amp30E17-*mCherry*	$HgCl_2$	2	180.67 ± 25.70	1.90 ± 0.27	6.2	4	Yes	37 ℃, 12 h	
J115-P*merT*-Amp32E17-*mCherry*	$HgCl_2$	> 10	-	-	6.2	4	Yes	37 ℃, 12 h	Fig. 5.12
J115-P*merT*-Amp33E17-*mCherry*	$HgCl_2$	2	92.33 ± 39.55	6.30 ± 2.70	6.2	4	Yes	37 ℃, 12 h	
J115-P*merT*-Amp30E22-*mCherry*	$HgCl_2$	2	669.67 ± 89.24	3.06 ± 0.41	6.2	4	Yes	37 ℃, 12 h	
J115-P*merT*-Amp32E22-*mCherry*	$HgCl_2$	10	3688.00 ± 810.99	34.68 ± 7.63	6.2	4	Yes	37 ℃, 12 h	
J115-P*merT*-Amp33E22-*mCherry*	$HgCl_2$	1	100.67 ± 41.88	7.19 ± 2.99	6.2	4	Yes	37 ℃, 12 h	

J115-P$_{merT}$-Amp33E16-*mCherry*	$HgCl_2$	10	411.00 ± 127.71	6.04 ± 1.88	6.2	4	Yes	37 °C, 12 h	
J115-P$_{merT}$-Amp33E20-*mCherry*	$HgCl_2$	2	263.33 ± 86.35	3.45 ± 1.13	6.2	4	Yes	37 °C, 12 h	
J115-P$_{merT}$-*lacZ*	$HgCl_2$	1	-	-	9.7 (with 0.02 - 0.04 % X-gal)	2 (paper-based)	Yes	37 °C, 90 - 150 min	Figs. 5.14f, 5.16a
J115-P$_{merT}$-*lacZ*	$HgCl_2$	10	-	-	9.7 (with 0.03 - 0.06 % CPRG)	2 (paper-based)	Yes	37 °C, 30 min	Fig. 5.15
J115-P$_{merT}$-*lacZ*	$HgCl_2$	1	-	-	9.7 (with 0.03 % CPRG)	2 (paper-based)	Yes	37 °C, 90 min	Fig. 5.15

*: LOD here is the target concentration used in each experiment that lead to significant sensor response.

**: Output at the LOD concentration induction.

***: Induction fold at the LOD concentration induction.

Appendix 5. Publications

Part of the **Chapter 1. Introduction** has been re-written and submitted to '*Handbook on Cell Biosensors*' (Springer) as Chapter 'Engineering prokaryote synthetic biology biosensors', and is currently in press.

Results of **Chapter 3. Signal Amplifying Gene Circuits Enable Ultrasensitive Cellular Sensors** and part of **Chapter 4. Designing a Cell-based Sensor Array for Field Testing** have been submitted as a journal article to *Nature Chemical Biology* (in press).

I have contributed as a co-author to the following journal article: Liu Q, Schumacher J, **Wan X**, Lou C and Wang B, "Orthogonality and burdens of heterologous AND gate gene circuits in *E. coli*", *ACS Synthetic Biology*, 2018, 7(2): 553-564.

References

Abil Z, Ellefson JW, Gollihar JD, Watkins E, Ellington AD, (2017) Compartmentalized partnered replication for the directed evolution of genetic parts and circuits. *Nature Protocols*, 12:2493–2512.

Adamala KP, Martin-Alarcon DA, Guthrie-Honea KR, Boyden ES, (2017) Engineering genetic circuit interactions within and between synthetic minimal cells. *Nature Chemistry*, 9:431–439.

Albert H, Dale EC, Lee E, Ow DW, (1995) Site-specific integration of DNA into wild-type and mutant lox sites placed in the plant genome. *The Plant Journal*, 7:649–659.

Alon U, (2007) An Introduction to Systems Biology: Design Principles of Biological Circuits. Chapman & Hall/CRC

Amaro F, Turkewitz AP, Martín-González A, Gutiérrez JC, (2011) Whole-cell biosensors for detection of heavy metal ions in environmental samples based on metallothionein promoters from Tetrahymena thermophila. *Microbial Biotechnology*, 4:513–522.

Andersen JB, Sternberg C, Poulsen LK, Bjørn SP, Givskov M, Molin S, (1998) New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Applied and Environmental Microbiology*, 64:2240–2246.

Anderson JC, Clarke EJ, Arkin AP, Voigt CA, (2006) Environmentally controlled invasion of cancer cells by engineered bacteria. *Journal of Molecular Biology*, 355:619–627.

Anderson JC, Voigt CA, Arkin AP, (2007) Environmental signal integration by a modular and gate. *Molecular Systems Biology*, 3:133.

Andres J, Bertin PN, (2016) The microbial genomics of arsenic. *FEMS Microbiology Reviews*, 40:299–322.

Ang J, Harris E, Hussey BJ, Kil R, McMillen DR, (2013) Tuning response curves for synthetic biology. *ACS Synthetic Biology*, 2:547–567.

Archer EJ, Robinson AB, Süel GM, (2012) Engineered E. coli that detect and respond to gut inflammation through nitric oxide sensing. *ACS Synthetic Biology*, 1:451–7.

Armbruster DA, Pry T, (2008) Limit of blank, limit of detection and limit of quantitation. *The Clinical Biochemist Reviews / Australian Association of Clinical Biochemists*, 29 Suppl 1:S49-52.

Arpino JAJ, Hancock EJ, Anderson J, Barahona M, Stan GB V., Papachristodoulou A, Polizzi K, (2013) Tuning the dials of synthetic biology. *Microbiology (United Kingdom)*, 159:1236–1253.

Barkay T, Miller SM, Summers AO, (2003) Bacterial mercury resistance from atoms to ecosystems. *FEMS Microbiology Reviews*, 27:355–384.

Barrios H, Valderrama B, Morett E, (1999) Compilation and analysis of sigma(54)-dependent promoter sequences. *Nucleic Acids Research*, 27:4305–4313.

Bashor CJ, Collins JJ, (2018) Understanding Biological Regulation Through Synthetic Biology. *Annual Review of Biophysics*, 47:399–423.

BBC News, (1986) 1986: Coal mine canaries made redundant. http://news.bbc.co.uk/onthisday/hi/dates/stories/december/30/newsid_2547000/2547587.stm. Accessed 27 Dec 2017

Belkin S, Yagur-Kroll S, Kabessa Y, Korouma V, Septon T, Anati Y, Zohar-Perez C, Rabinovitz Z, Nussinovitch A, Agranat AJ, (2017) Remote detection of buried landmines using a bacterial sensor. *Nature Biotechnology*, 35:308–310.

Bergkessel M, Guthrie C, (2013) Colony PCR. *Methods in Enzymology*, 529:299–309.

Bernard E, Wang B, (2017) Synthetic cell-based sensors with programmed selectivity and sensitivity. In: Methods in Molecular Biology. pp 349–363

Bernhoft RA, (2012) Mercury toxicity and treatment: A review of the literature. *Journal of Environmental and Public Health*, 2012:460508.

Berset Y, Merulla D, Joublin A, Hatzimanikatis V, van der Meer JR, (2017) Mechanistic Modeling of Genetic Circuits for ArsR Arsenic Regulation. *ACS Synthetic Biology*, 6:862–874.

Bhattacharjee H, Rosen BP, (2007) Arsenic Metabolism in Prokaryotic and Eukaryotic Microbes. *Molecular Microbiology of Heavy Metals*, 6:371–406.

Biran I, Rissin DM, Ron EZ, Walt DR, (2003) Optical imaging fiber-based live bacterial cell array biosensor. *Analytical Biochemistry*, 315:106–113.

Bonnet J, Subsoontorn P, Endy D, (2012) Rewritable digital data storage in live cells via engineered control of recombination directionality. *Proceedings of the National Academy of Sciences*, 109:8884–8889.

Bonnet J, Yin P, Ortiz ME, Subsoontorn P, Endy D, (2013) Amplifying genetic logic gates. *Science (New York, NY)*, 340:599–603.

Bourdeau RW, Lee-Gosselin A, Lakshmanan A, Farhadi A, Kumar SR, Nety SP, Shapiro MG, (2018) Acoustic reporter genes for noninvasive imaging of microorganisms in mammalian hosts. *Nature*, 553:86–90.

Bradley RW, Buck M, Wang B, (2016a) Recognizing and engineering digital-like logic gates and switches in gene regulatory networks. *Current Opinion in Microbiology*, 33:74–82.

Bradley RW, Buck M, Wang B, (2016b) Tools and Principles for Microbial Gene Circuit Engineering. *Journal of Molecular Biology*, 428:862–888.

Bradley RW, Wang B, (2015) Designer cell signal processing circuits for biotechnology. *New Biotechnology*, 32:635–643.

Brewster RC, Jones DL, Phillips R, (2012) Tuning Promoter Strength through RNA Polymerase Binding Site Design in Escherichia coli. *PLoS Computational Biology*, 8:e1002811.

Brophy JAN, Voigt CA, (2016) Antisense transcription as a tool to tune gene expression. *Molecular Systems Biology*, 12:854–854.

Brophy JAN, Voigt CA, (2014) Principles of genetic circuit design. *Nature Methods*, 11:508–520.

Buffi N, Merulla D, Beutier J, Barbaud F, Beggah S, Van Lintel H, Renaud P, Roelof van der Meer J, (2011) Development of a microfluidics biosensor for agarose-bead immobilized Escherichia coli bioreporter cells for arsenite detection in aqueous samples. *Lab on a Chip*, 11:2369–2377.

Bulich AA, Isenberg DL, (1981) Use of the luminescent bacterial system for the rapid assessment of aquatic toxicity. *ISA Transactions*, 20:29–33.

Callura JM, Dwyer DJ, Isaacs FJ, Cantor CR, Collins JJ, (2010) Physiology Using Synthetic Riboregulators. *Proceedings of the National Academy of Sciences*, 107:15898–15903.

Cameron DE, Bashor CJ, Collins JJ, (2014) A brief history of synthetic biology. *Nature Reviews Microbiology*, 12:381–90.

Cameron DE, Collins JJ, (2014) Tunable protein degradation in bacteria. *Nature Biotechnology*, 32:1–8.

Cao Y, Feng Y, Ryser MD, Zhu K, Herschlag G, Cao C, Marusak K, Zauscher S, You L, (2017) Programmable assembly of pressure sensors using pattern-forming bacteria. *Nature Biotechnology*, 35:1087–1093.

Carrilho E, Martinez AW, Whitesides GM, (2009) Understanding wax printing: A simple micropatterning process for paper-based microfluidics. *Analytical Chemistry*, 81:7091–7095.

Caschera F, Noireaux V, (2014) Synthesis of 2.3 mg/ml of protein with an all Escherichia coli cell-free transcription-translation system. *Biochimie*, 99:162–168.

Cayron J, Prudent E, Escoffier C, Gueguen E, Mandrand-Berthelot MA, Pignol D, Garcia D, Rodrigue A, (2017) Pushing the limits of nickel detection to nanomolar range using a set of engineered bioluminescent Escherichia coli. *Environmental Science and Pollution Research*, 24:4–14.

Cerminati S, Soncini FC, Checa SK, (2015) A sensitive whole-cell biosensor for the simultaneous detection of a broad-spectrum of toxic heavy metal ions. *Chemical Communications*, 51:5917–5920.

Cerminati S, Soncini FC, Checa SK, (2011) Selective detection of gold using genetically engineered bacterial reporters. *Biotechnology and Bioengineering*, 108:2553–2560.

Cervantes C, Ji G, Ramirez J, Silver S, (1994) Resistance to arsenic compounds in microorganisms.

FEMS Microbiology Reviews, 15:355–367.

Cevenini L, Lopreside A, Calabretta MM, D'Elia M, Simoni P, Michelini E, Roda A, (2018) A novel bioluminescent NanoLuc yeast-estrogen screen biosensor (nanoYES) with a compact wireless camera for effect-based detection of endocrine-disrupting chemicals. *Analytical and Bioanalytical Chemistry*, 410:1237–1246.

Chang CC, Lin LY, Zou XW, Huang CC, Chan NL, (2015) Structural basis of the mercury(II)-mediated conformational switching of the dual-function transcriptional regulator MerR. *Nucleic Acids Research*, 43:7612–7623.

Chang TMS, Prakash S, (2001) Procedures for microencapsulation of enzymes, cells and genetically engineered microorganisms. *Applied Biochemistry and Biotechnology - Part B Molecular Biotechnology*, 17:249–260.

Chappell J, Jensen K, Freemont PS, (2013) Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology. *Nucleic Acids Research*, 41:3471–3481.

Chen J, Rosen BP, (2014) Biosensors for inorganic and organic arsenicals. *Biosensors*, 4:494–512.

Chen Y, Ho JML, Shis DL, Gupta C, Long J, Wagner DS, Ott W, Josić K, Bennett MR, (2018) Tuning the dynamic range of bacterial promoters regulated by ligand-inducible transcription factors. *Nature Communications*, 9:64.

Cheng AA, Lu TK, (2012) Synthetic Biology: An Emerging Engineering Discipline. *Annual Review of Biomedical Engineering*, 14:155–178.

Choi PJ, Cai L, Frieda K, Xie XS, (2008) A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science*, 322:442–446.

Chowdhury SR, Yanful EK, Pratt AR, (2011) Arsenic removal from aqueous solutions by mixed magnetite-maghemite nanoparticles. *Environmental Earth Sciences*, 64:411–423.

Church GM, Elowitz MB, Smolke CD, Voigt C a, Weiss R, (2014) Realizing the potential of synthetic biology. *Nature Reviews Molecular Cell Biology*, 15:289–94.

Courbet A, Endy D, Renard E, Molina F, Bonnet J, (2015) Detection of pathological biomarkers in human clinical samples via amplifying genetic switches and logic gates. *Science Translational Medicine*, 7:289ra83.

Cox RS, Surette MG, Elowitz MB, (2007) Programming gene expression with combinatorial promoters. *Molecular Systems Biology*, 3:145.

Daeffler KN, Galley JD, Sheth RU, Ortiz-Velez LC, Bibb CO, Shroyer NF, Britton RA, Tabor JJ, (2017) Engineering bacterial thiosulfate and tetrathionate sensors for detecting gut inflammation. *Molecular Systems Biology*, 13:923.

Dana G V., Kuiken T, Rejeski D, Snow AA, (2012) Synthetic biology: Four steps to avoid a synthetic-biology disaster. *Nature*, 483:29–29.

Danino T, Prindle A, Kwong GA, Skalak M, Li H, Allen K, Hasty J, Bhatia SN, (2015) Programmable probiotics for non-invasive urinary detection of cancer. *Science Translational Medicine*, 7:1–36.

Das J, Sarkar P, (2016) A new dipstick colorimetric sensor for detection of arsenate in drinking water. *Environ Sci: Water Res Technol*, 2:693–704.

Date A, Pasini P, Daunert S, (2007) Construction of spores for portable bacterial whole-cell biosensing systems. *Analytical Chemistry*, 79:9391–9397.

Daunert S, Barrett G, Feliciano JS, Shetty RS, Shrestha S, Smith-Spencer W, (2000) Genetically engineered whole-cell sensing systems: coupling biological recognition with reporter genes. *Chemical Reviews*, 100:2705–2738.

De Los Santos ELC, Meyerowitz JT, Mayo SL, Murray RM, (2016) Engineering Transcriptional Regulator Effector Specificity Using Computational Design and in Vitro Rapid Prototyping: Developing a Vanillin Sensor. *ACS Synthetic Biology*, 5:287–295.

De Mora K, Joshi N, Balint BL, Ward FB, Elfick A, French CE, (2011) A pH-based biosensor for detection of arsenic in drinking water. *Analytical and Bioanalytical Chemistry*, 400:1031–1039.

Dhuldhaj UP, Sharma NK, Singh S, (2012) Microbial Removal of Arsenic: An Overview. In: Bioremediation of Pollutants

Didovyk A, Tonooka T, Tsimring L, Hasty J, (2017) Rapid and Scalable Preparation of Bacterial Lysates for Cell-Free Gene Expression. *ACS Synthetic Biology*, 6:2198–2208.

Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS, (1991) "Touchdown" PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research*, 19:4008.

Duan F, March JC, (2010) Engineered bacterial communication prevents Vibrio cholerae virulence in an infant mouse model. *Proceedings of the National Academy of Sciences of the United States of America*, 107:11260–4.

Duyen TTM, Matsuura H, Ujiie K, Muraoka M, Harada K, Hirata K, (2017) Paper-based colorimetric biosensor for antibiotics inhibiting bacterial protein synthesis. *Journal of Bioscience and Bioengineering*, 123:96–100.

Egbert RG, Klavins E, (2012) Fine-tuning gene networks using simple sequence repeats. *Proceedings of the National Academy of Sciences of the United States of America*, 109:16817–22.

Ellis T, Wang X, Collins JJ, (2009) Diversity-Based, Model-Guided Construction of Synthetic Gene Networks with Predicted Functions. *Nature Biotechnology*, 27:465–471.

Endy D, (2005) Foundations for engineering biology. *Nature*, 438:449–453.

Esvelt KM, Carlson JC, Liu DR, (2011) A system for the continuous directed evolution of biomolecules. *Nature*, 472:499–503.

Fernandez-Rodriguez J, Moser F, Song M, Voigt CA, (2017) Engineering RGB color vision into Escherichia coli. *Nature Chemical Biology*, 13:706–708.

Fernandez-Rodriguez J, Voigt CA, (2016) Post-translational control of genetic circuits using Potyvirus proteases. *Nucleic Acids Research*, 44:6493–6502.

Flanagan S, Johnston R, Zheng Y, (2012) Arsenic in tube well water in Bangladesh: health and economic impacts and implications for arsenic mitigation. *Bulletin of the World Health Organization*, 90:839–846.

Folliard T, Steel H, Prescott TP, Wadhams G, Rothschild LJ, Papachristodoulou A, (2017) A Synthetic Recombinase-Based Feedback Loop Results in Robust Expression. *ACS Synthetic Biology*, 6:1663–1671.

Forbes NS, (2010) Engineering the perfect (bacterial) cancer therapy. *Nature Reviews Cancer*, 10:785–794.

Friedland AE, Lu TK, Wang X, Shi D, Church G, Collins JJ, (2009) Synthetic gene networks that count. *Science (New York, NY)*, 324:1199–202.

Fritz BR, Jamil OK, Jewett MC, (2015) Implications of macromolecular crowding and reducing conditions for in vitro ribosome construction. *Nucleic Acids Research*, 43:4774–4784.

Fujimoto H, Wakabayashi M, Yamashiro H, Maeda I, Isoda K, Kondoh M, Kawase M, Miyasaka H, Yagi K, (2006) Whole-cell arsenite biosensor using photosynthetic bacterium Rhodovulum sulfidophilum: Rhodovulum sulfidophilum as an arsenite biosensor. *Applied Microbiology and Biotechnology*, 73:332–338.

Garamella J, Marshall R, Rustad M, Noireaux V, (2016) The All E. coli TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology. *ACS Synthetic Biology*, 5:344–355.

Gardner TS, Cantor CR, Collins JJ, (2000) Construction of a genetic toggle switch in Escherichia coli. *Nature*, 403:339–342.

Ge X, Luo D, Xu J, (2011) Cell-free protein expression under macromolecular crowding conditions. *PLoS ONE*, 6:e28707.

Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison C a, Smith HO, (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, 6:343–5.

Gregor C, Gwosch KC, Sahl SJ, Hell SW, (2018) Strongly enhanced bacterial bioluminescence with the ilux operon for single-cell imaging. *Proceedings of the National Academy of Sciences*, 115:962–967.

Gronbach K, Eberle U, Müller M, Olschläger T a, Dobrindt U, Leithäuser F, Niess JH, Döring G, Reimann J, Autenrieth IB, Frick J-S, (2010) Safety of probiotic Escherichia coli strain Nissle 1917 depends on intestinal microbiota and adaptive immunity of the host. *Infection and Immunity*, 78:3036–46.

Gui Q, Lawson T, Shan S, Yan L, Liu Y, (2017) The Application of Whole Cell-Based Biosensors for Use in Environmental Analysis and in Medical Diagnostics. *Sensors*, 17:1623.

Gupta S, Bram EE, Weiss R, (2013) Genetically programmable pathogen sense and destroy. *ACS Synthetic Biology*, 2:715–23.

Guzina J, Djordjevic M, (2017) Mix-and-matching as a promoter recognition mechanism by ECF σ factors. *BMC Evolutionary Biology*, 17:1–14.

Gyorgy A, Jiménez JI, Yazbek J, Huang HH, Chung H, Weiss R, Del Vecchio D, (2015) Isocost lines describe the cellular economy of genetic circuits. *Biophysical Journal*, 109:639–646.

Hakkila KM, Nikander PA, Junttila SM, Lamminmäki UJ, Virta MP, (2011) Cd-specific mutants of mercury-sensing regulatory protein MerR, generated by directed evolution. *Applied and Environmental Microbiology*, 77:6215–6224.

Halleran AD, Murray RM, (2018) Cell-Free and in Vivo Characterization of Lux, Las, and Rpa Quorum Activation Systems in E. coli. *ACS Synthetic Biology*, 7:752–755.

Hao N, Krishna S, Ahlgren-Berg A, Cutts EE, Shearwin KE, Dodd IB, (2014) Road rules for traffic on DNA - Systematic analysis of transcriptional roadblocking in vivo. *Nucleic Acids Research*, 42:8861–8872.

Hauschild KE, Metzler RE, Arndt H-D, Moretti R, Raffaelle M, Dervan PB, Ansari AZ, (2005) Temperature-sensitive protein-DNA dimerizers. *Proceedings of the National Academy of Sciences of the United States of America*, 102:5008–5013.

Hindmarsh JT, McCurdy RF, (1986) Clinical and environmental aspects of arsenic toxicity. *Critical Reviews in Clinical Laboratory Sciences*, 23:315–47.

Ho CL, Tan HQ, Chua KJ, Kang A, Lim KH, Ling KL, Yew WS, Lee YS, Thiery JP, Chang MW, (2018) Engineered commensal microbes for diet-mediated colorectal-cancer chemoprevention. *Nature Biomedical Engineering*, 2:27–37.

Hong SH, Ntai I, Haimovich AD, Kelleher NL, Isaacs FJ, Jewett MC, (2014) Cell-free protein synthesis from a release factor 1 deficient escherichia coli activates efficient and multiple site-specific nonstandard amino acid incorporation. *ACS Synthetic Biology*, 3:398–409.

Hooshangi S, Thiberge S, Weiss R, (2005) Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proceedings of the National Academy of Sciences of the United States of America*, 102:3581–6.

Huang CW, Wei CC, Liao VHC, (2015a) A low cost color-based bacterial biosensor for measuring arsenic in groundwater. *Chemosphere*, 141:44–49.

Huang CW, Yang SH, Sun MW, Liao VHC, (2015b) Development of a set of bacterial biosensors for simultaneously detecting arsenic and mercury in groundwater. *Environmental Science and Pollution Research*, 22:10206–10213.

Hwang IY, Koh E, Kim HR, Yew WS, Chang MW, (2016) Reprogrammable microbial cell-based therapeutics against antibiotic-resistant bacteria. *Drug Resistance Updates*, 27:59–71.

Hwang IY, Tan MH, Koh E, Ho CL, Poh CL, Chang MW, (2014) Reprogramming microbes to be pathogen-seeking killers. *ACS Synthetic Biology*, 3:228–37.

Hynninen A, Tõnismann K, Virta M, (2010) Improving the sensitivity of bacterial bioreporters for heavy metals. *Bioengineered Bugs*, 1:132–138.

Inniss MC, Silver PA, (2013) Building synthetic memory. *Current Biology*, 23:R812–R816.

Jiang F, Doudna JA, (2017) CRISPR–Cas9 Structures and Mechanisms. *Annual Review of Biophysics*, 46:505–529.

Joe MH, Lee KH, Lim SY, Im SH, Song HP, Lee IS, Kim DH, (2012) Pigment-based whole-cell biosensor system for cadmium detection using genetically engineered Deinococcus radiodurans. *Bioprocess and Biosystems Engineering*, 35:265–272.

Joshi N, Wang X, Montgomery L, Elfick A, French CE, (2009) Novel approaches to biosensors for detection of arsenic in drinking water. *Desalination*, 248:517–523.

Jovanovic M, James EH, Burrows PC, Rego FGM, Buck M, Schumacher J, (2011) Regulation of the co-evolved HrpR and HrpS AAA+ proteins required for Pseudomonas syringae pathogenicity. *Nature Communications*, 2:177.

Karig D, Weiss R, (2005) Signal-amplifying genetic circuit enables in vivo observation of weak promoter activation in the Rhl quorum sensing system. *Biotechnology and Bioengineering*, 89:709–718.

Karim AS, Jewett MC, (2016) A cell-free framework for rapid biosynthetic pathway prototyping and enzyme discovery. *Metabolic Engineering*, 36:116–126.

Kaur H, Kumar R, Babu JN, Mittal S, (2015) Advances in arsenic biosensor development--a comprehensive review. *Biosensors & Bioelectronics*, 63:533–45.

Kawano M, Aravind L, Storz G, (2007) An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. *Molecular Microbiology*, 64:738–754.

Kearns J, Tyson J, (2012) Improving the accuracy and precision of an arsenic field test kit: increased reaction time and digital image analysis. *Analytical Methods*, 4:1693–1698.

Khatun MA, Hoque MA, Zhang Y, Lu T, Cui L, Zhou N-Y, Feng Y, (2018) Bacterial Consortium-Based Sensing System for Detecting Organophosphorus Pesticides. *Analytical Chemistry*, acs.analchem.8b02709.

Kigawa T, Yabuki T, Matsuda N, Matsuda T, Nakajima R, Tanaka A, Yokoyama S, (2004) Preparation of Escherichia coli cell extract for highly productive cell-free protein expression. *Journal of Structural and Functional Genomics*, 5:63–68.

Kim DM, Kigawa T, Choi CY, Yokoyama S, (1996) A highly efficient cell-free protein synthesis system from Escherichia coli. *European Journal of Biochemistry / FEBS*, 239:881–886.

Kim HJ, Jeong H, Lee SJ, (2018) Synthetic biology for microbial heavy metal biosensors. *Analytical and Bioanalytical Chemistry*, 410:1191–1203.

Kim HJ, Lim JW, Jeong H, Lee SJ, Lee DW, Kim T, Lee SJ, (2016) Development of a highly specific and sensitive cadmium and lead microbial biosensor using synthetic CadC-T7 genetic circuitry. *Biosensors and Bioelectronics*, 79:701–708.

King JMH, DiGrazia PM, Applegate B, Burlage R, Sanseverino J, Dunbar P, Larimer F, Sayler GS, (1990) Rapid, Sensitive Bioluminescent Reporter Technology for Naphthalene Exposure and Biodegradation. *Science*, 249:778–781.

Kosuri S, Goodman DB, Cambray G, Mutalik VK, Gao Y, Arkin AP, Endy D, Church GM, (2013) Composability of regulatory sequences controlling transcription and translation in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 110:14024–9.

Kotula JW, Kerns SJ, Shaket LA, Siraj L, Collins JJ, Way JC, Silver PA, (2014) Programmable bacteria detect and record an environmental signal in the mammalian gut. *Proceedings of the National Academy of Sciences*, 111:4838–4843.

Kurland T, Faro SN, Siedler H, (1960) Minamata disease. The outbreak of a neurologic disorder in Minamata, Japan, and its relationship to the ingestion of seafood contaminated by mercuric compounds. *World Neurology*, 1:370–395.

Kuruma Y, Ueda T, (2015) The PURE system for the cell-free synthesis of membrane proteins. *Nature Protocols*, 10:1328–1344.

Kuznetsova IM, Turoverov KK, Uversky VN, (2014) What macromolecular crowding can do to a protein. *International Journal of Molecular Sciences*, 15:23090–23140.

Kwon YC, Jewett MC, (2015) High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Scientific Reports*, 5:1–8.

Landry BP, Palanki R, Dyulgyarov N, Hartsough LA, Tabor JJ, (2018) Phosphatase activity tunes two-component system sensor detection threshold. *Nature Communications*, 9:1433.

Lasa I, Toledo-Arana A, Dobin A, Villanueva M, de los Mozos IR, Vergara-Irigaray M, Segura V, Fagegaltier D, Penades JR, Valle J, Solano C, Gingeras TR, (2011) Genome-wide antisense transcription drives mRNA processing in bacteria. *Proceedings of the National Academy of Sciences*, 108:20172–20177.

Li J, Gu L, Aach J, Church GM, (2014) Improved cell-free RNA and protein synthesis system. *PLoS ONE*, 9:e106232.

Li L, Liang J, Hong W, Zhao Y, Sun S, Yang X, Xu A, Hang H, Wu L, Chen S, (2015) Evolved bacterial biosensor for arsenite detection in environmental water. *Environmental Science and Technology*,

49:6149–6155.

Lin HY, Bledsoe PJ, Stewart V, (2007) Activation of yeaR-yoaG operon transcription by the nitrate-responsive regulator NarL is independent of oxygen-responsive regulator Fnr in Escherichia coli K-12. *Journal of Bacteriology*, 189:7539–7548.

Liu Q, Schumacher J, Wan X, Lou C, Wang B, (2018a) Orthogonality and Burdens of Heterologous and Gate Gene Circuits in E. coli. *ACS Synthetic Biology*, 7:553–564.

Liu X, Yuk H, Lin S, Parada GA, Tang TC, Tham E, de la Fuente-Nunez C, Lu TK, Zhao X, (2018b) 3D Printing of Living Responsive Materials and Devices. *Advanced Materials*, 30:107821.

Loening AM, Wu AM, Gambhir SS, (2007) Red-shifted Renilla reniformis luciferase variants for imaging in living subjects. *Nature Methods*, 4:641–643.

Ma D, Shen L, Wu K, Diehnelt CW, Green AA, (2018) Low-cost detection of norovirus using paper-based cell-free systems and synbody-based viral enrichment. *Synthetic Biology*, 3:ysy018.

Ma KC, Perli SD, Lu TK, (2016) Foundations and Emerging Paradigms for Computing in Living Cells. *Journal of Molecular Biology*, 428:893–915.

Mahr R, Frunzke J, (2016) Transcription factor-based biosensors in biotechnology: current state and future prospects. *Applied Microbiology and Biotechnology*, 100:79–90.

Mannan AA, Liu D, Zhang F, Oyarzún DA, (2017) Fundamental Design Principles for Transcription-Factor-Based Metabolite Biosensors. *ACS Synthetic Biology*, 6:1851–1859.

Marshall R, Maxwell CS, Collins SP, Jacobsen T, Luo ML, Begemann MB, Gray BN, January E, Singer A, He Y, Beisel CL, Noireaux V, (2018) Rapid and Scalable Characterization of CRISPR Technologies Using an E. coli Cell-Free Transcription-Translation System. *Molecular Cell*, 69:146–157.e3.

Mary P Hall, James Unch, Brock F Binkowski, Michael P Valley, Braeden L Butler, Monika G Wood, Paul Otto, Kristopher Zimmerman, Gediminas Vidugiris, Thomas Machleidt, Matthew P Robers, Hélène A Benink, Christopher T Eggers, Michael R Slater, Poncho L Meisenheimer, Dieter H Klauber, Frank Fan, Lance P Encell, Keith V Wood, (2012) Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. ACS Chem Biol

Mathema VB, Thakuri BC, Sillanpää M, (2011) Bacterial mer operon-mediated detoxification of mercurial compounds: A short review. *Archives of Microbiology*, 193:837–844.

Maxwell CS, Jacobsen T, Marshall R, Noireaux V, Beisel CL, (2018) A detailed cell-free transcription-translation-based assay to decipher CRISPR protospacer-adjacent motifs. *Methods*, 143:48–57.

Merrin J, Kumar P, Libchaber A, (2011) Effects of pressure and temperature on the binding of RecA protein to single-stranded DNA. *Proceedings of the National Academy of Sciences*, 108:19913–19918.

Merulla D, Hatzimanikatis V, van der Meer JR, (2013) Tunable reporter signal production in feedback-uncoupled arsenic bioreporters. *Microbial Biotechnology*, 6:503–514.

Merulla D, van der Meer JR, (2016) Regulatable and modulable background expression control in prokaryotic synthetic circuits by auxiliary repressor binding sites. *ACS Synthetic Biology*, 5:36–45.

Michelini E, Cevenini L, Mezzanotte L, Leskinen P, Virta M, Karp M, Roda A, (2008) A sensitive recombinant cell-based bioluminescent assay for detection of androgen-like compounds. Nat Protoc. doi: 10.1038/nprot.2008.189

Miller WG, Brandl MT, Quiñones B, Lindow SE, (2001) Biological Sensor for Sucrose Availability: Relative Sensitivities of Various Reporter Genes. *Applied and Environmental Microbiology*, 67:1308–1317.

Mimee M, Nadeau P, Hayward A, Carim S, Flanagan S, Jerger L, Collins J, McDonnell S, Swartwout R, Citorik RJ, Bulović V, Langer R, Traverso G, Chandrakasan AP, Lu TK, (2018) An ingestible bacterial-electronic system to monitor gastrointestinal health. *Science*, 360:915–918.

Mimee M, Tucker AC, Voigt CA, Lu TK, (2015) Programming a Human Commensal Bacterium, Bacteroides thetaiotaomicron, to Sense and Respond to Stimuli in the Murine Gut Microbiota. *Cell Systems*, 1:62–71.

Misra TK, Brown NL, Haberstroh L, Schmidt A, Goddette D, Silver S, (1985) Mercuric reductase structural genes from plasmid R100 and transposon Tn501: functional domains of the enzyme.

Gene, 34:253–262.

Möckli N, Auerbach D, (2004) Quantitative β-galactosidase assay suitable for high-throughput applications in the yeast two-hybrid system. *BioTechniques*, 36:872–876.

Moe-Behrens GHG, Davis R, Haynes KA, (2013) Preparing synthetic biology for the world. *Frontiers in Microbiology*, 4:5.

Moon TS, Clarke EJ, Groban ES, Tamsir A, Clark RM, Eames M, Kortemme T, Voigt CA, (2011) Construction of a genetic multiplexer to toggle between chemosensory pathways in Escherichia coli. *Journal of Molecular Biology*, 406:215–227.

Moon TS, Lou C, Tamsir A, Stanton BC, Voigt CA, (2012) Genetic programs constructed from layered logic gates in single cells. *Nature*, 491:249–253.

Moore B, (1960) A NEW SCREEN TEST AND SELECTIVE MEDIUM FOR THE RAPID DETECTION OF EPIDEMIC STRAINS OF STAPH. AUREUS. *The Lancet*, 276:453–458.

Mutalik VK, Guimaraes JC, Cambray G, Lam C, Christoffersen MJ, Mai QA, Tran AB, Paull M, Keasling JD, Arkin AP, Endy D, (2013) Precise and reliable gene expression via standard transcription and translation initiation elements. *Nature Methods*, 10:354–360.

Nassif N, Bouvet O, Noelle Rager M, Roux C, Coradin T, Livage J, (2002) Living bacteria in silica gels. *Nature Materials*, 1:42–44.

Nealson KH, Belz A, McKee B, (2002) Breathing metals as a way of life: Geobiology in action. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, 81:215–222.

Nelson MD, Fitch DHA, (2011) Overlap Extension PCR: An Efficient Method for Transgene Construction. In: Molecular Methods for Evolutionary Genetics. pp 459–470

Niederholtmeyer H, Sun ZZ, Hori Y, Yeung E, Verpoorte A, Murray RM, Maerkl SJ, (2015) Rapid cell-free forward engineering of novel genetic ring oscillators. *ELife*, 4:1–18.

Nielsen AAK, Der BS, Shin J, Vaidyanathan P, Paralanov V, Strychalski EA, Ross D, Densmore D, Voigt CA, (2016) Genetic circuit design automation. *Science*, 352:aac7341.

Nistala GJ, Wu K, Rao C V., Bhalerao KD, (2010) A modular positive feedback-based gene amplifier. *Journal of Biological Engineering*, 4:4.

Oberdoerffer P, (2003) Unidirectional Cre-mediated genetic inversion in mice using the mutant loxP pair lox66/lox71. *Nucleic Acids Research*, 31:140e–140.

Ochman H, Gerber AS, Hartl DL, (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics*, 120:621–623.

Oda Y, Nakamura S ichi, Oki I, Kato T, Shinagawa H, (1985) Evaluation of the new system (umu-test) for the detection of environmental mutagens and carcinogens. Mutat Res Mutagen Relat Subj. doi: 10.1016/0165-1161(85)90062-7

Ohi G, Seki H, Minowa K, Ohsawa M, Mizoguchi I, Sugimori F, (1981) Lead pollution in Tokyo - The pigeon reflects its amelioration. *Environmental Research*, 26:125–129.

Olivieri G, Novakovic M, Savaskan E, Meier F, Baysang G, Brockhaus M, Müller-Spahn F, (2002) The effects of β-estradiol on SHSY5Y neuroblastoma cells during heavy metal induced oxidative stress, neurotoxicity and β-amyloid secretion. *Neuroscience*, 113:849–855.

Olorunniji FJ, Rosser SJ, Stark WM, (2016) Site-specific recombinases: molecular machines for the Genetic Revolution. *Biochemical Journal*, 473:673–684.

Olson EJ, Hartsough LA, Landry BP, Shroff R, Tabor JJ, (2014) Characterizing bacterial gene circuit dynamics with optically programmed gene expression signals. *Nature Methods*, 11:449–55.

Papi RM, Chaitidou SA, Trikka FA, Kyriakidis DA, (2005) Encapsulated Escherichia coli in alginate beads capable of secreting a heterologous pectin lyase. *Microbial Cell Factories*, 4:35.

Pardee K, Green A a, Ferrante T, Cameron DE, Daleykeyser A, Yin P, (2014) Paper-Based Synthetic Gene Networks. *Cell*, 159:940–954.

Pardee K, Green AA, Takahashi MK, Braff D, Lambert G, Lee JW, Ferrante T, Ma D, Donghia N, Fan M, Daringer NM, Bosch I, Dudley DM, O'Connor DH, Gehrke L, Collins JJ, (2016a) Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell*, 165:1255–1266.

Pardee K, Slomovic S, Nguyen PQ, Lee JW, Donghia N, Burrill D, Ferrante T, McSorley FR, Furuta

Y, Vernet A, Lewandowski M, Boddy CN, Joshi NS, Collins JJ, (2016b) Portable, On-Demand Biomolecular Manufacturing. *Cell*, 167:248–259.e12.

Park J-D, Zheng W, (2012) Human Exposure and Health Effects of Inorganic and Elemental Mercury. *Journal of Preventive Medicine & Public Health Prev Med Public Health*, 34445:344–352.

Pédelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS, (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nature Biotechnology*, 24:79–88.

Peking iGEM, (2013) Aromatics Scouts. http://2013.igem.org/Team:Peking/Project/BioSensors. Accessed 3 Jun 2018

Pelechano V, Steinmetz LM, (2013) Gene regulation by antisense transcription. *Nature Reviews Genetics*, 14:880–893.

Pellinen T, Huovinen T, Karp M, (2004) A cell-free biosensor for the detection of transcriptional inducers using firefly luciferase as a reporter. *Analytical Biochemistry*, 330:52–57.

Perez JG, Stark JC, Jewett MC, (2016) Cell-free synthetic biology: Engineering beyond the cell. *Cold Spring Harbor Perspectives in Biology*, 8:a023853.

Pinheiro VB, Taylor AI, Cozens C, Abramov M, Renders M, Zhang S, Chaput JC, Wengel J, Peak-Chew SY, McLaughlin SH, Herdewijn P, Holliger P, (2012) Synthetic genetic polymers capable of heredity and evolution. *Science*, 336:341–344.

Potvin-Trottier L, Lord ND, Vinnicombe G, Paulsson J, (2016) Synchronous long-term oscillations in a synthetic gene circuit. *Nature*, 538:514–517.

Pous N, Balaguer MD, Colprim J, Puig S, (2018) Opportunities for groundwater microbial electro-remediation. *Microbial Biotechnology*, 11:119–135.

Power B, Liu X, Germaine KJ, Ryan D, Brazil D, Dowling DN, (2011) Alginate beads as a storage, delivery and containment system for genetically modified PCB degrader and PCB biosensor derivatives of Pseudomonas fluorescens F113. *Journal of Applied Microbiology*, 110:1351–1358.

Prezioso VR, Jahns A, (2000) Using gradient PCR to determine the optimum annealing temperature. *Bioresearch Online*, 5–6.

Prindle A, Samayoa P, Razinkov I, Danino T, Tsimring LS, Hasty J, (2012) A sensing array of radically coupled genetic "biopixels." *Nature*, 481:39–44.

Purnick PEM, Weiss R, (2009) The second wave of synthetic biology: From modules to systems. *Nature Reviews Molecular Cell Biology*, 10:410–422.

Qian Y, Huang H-H, Jiménez JI, Del Vecchio D, (2017) Resource competition shapes the response of genetic circuits. *ACS Synthetic Biology*, 6:1263−1272.

Qin XJ, Hudson LG, Liu W, Ding W, Cooper KL, Liu KJ, (2008) Dual actions involved in arsenite-induced oxidative DNA damage. *Chemical Research in Toxicology*, 21:1806–1813.

Quiles-Puchalt N, Tormo-Más MÁ, Campoy S, Toledo-Arana A, Monedero V, Lasa Í, Novick RP, Christie GE, Penadés JR, (2013) A super-family of transcriptional activators regulates bacteriophage packaging and lysis in Gram-positive bacteria. *Nucleic Acids Research*, 41:7260–7275.

Quillardet P, Huisman O, D'Ari R, Hofnung M, (1982) SOS chromotest, a direct assay of induction of an SOS function in Escherichia coli K-12 to measure genotoxicity. Proc Natl Acad Sci. doi: 10.1073/pnas.79.19.5971

Rajkumar AS, Dénervaud N, Maerkl SJ, (2013) Mapping the fine structure of a eukaryotic promoter input-output function. *Nature Genetics*, 45:1207–1215.

Rampioni G, D'Angelo F, Messina M, Zennaro A, Kuruma Y, Tofani D, Leoni L, Stano P, (2018) Synthetic cells produce a quorum sensing chemical signal perceived by: Pseudomonas aeruginosa. *Chemical Communications*, 54:2090–2093.

Rampley CPN, Davison PA, Qian P, Preston GM, Hunter CN, Thompson IP, Wu LJ, Huang WE, (2017) Development of SimCells as a novel chassis for functional biosensors. *Scientific Reports*, 7:7261.

Rasmussen M, Minteer SD, (2015) Long-term arsenic monitoring with an Enterobacter cloacae microbial fuel cell. *Bioelectrochemistry*, 106:207–212.

Ratnaike RN, (2003) Acute and chronic arsenic toxicity. *Postgraduate Medical Journal*, 79:391–396.

Ravenscroft P, Brammer H, Richards K, (2009) Arsenic Pollution: A Global Synthesis

Registry of Standard Biological Parts, (2017) Ribosome Binding Sites/Prokaryotic/Constitutive/Community Collection. http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Community_Collection. Accessed 4 Jun 2018

Reid BJ, Sample KT, Macleod CJ, Weitz HJ, Paton GI, (1998) Feasibility of using prokaryote biosensors to assess acute toxicity of polycyclic aromatic hydrocarbons. *FEMS Microbiology Letters*, 169:227–233.

Rhodius VA, Segall-Shapiro TH, Sharon BD, Ghodasara A, Orlova E, Tabakh H, Burkhardt DH, Clancy K, Peterson TC, Gross C a, Voigt C a, (2013) Design of orthogonal genetic switches based on a crosstalk map of σs, anti-σs, and promoters. *Molecular Systems Biology*, 9:702.

Riglar DT, Giessen TW, Baym M, Kerns SJ, Niederhuber MJ, Bronson RT, Kotula JW, Gerber GK, Way JC, Silver PA, (2017) Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation. *Nature Biotechnology*, 35:653–658.

Roda A, Cevenini L, Michelini E, Branchini BR, (2011) A portable bioluminescence engineered cell-based biosensor for on-site applications. Biosens Bioelectron. doi: 10.1016/j.bios.2011.02.022

Roda A, Guardigli M, Michelini E, Mirasoli M, (2009) Bioluminescence in analytical chemistry and in vivo imaging. *TrAC - Trends in Analytical Chemistry*, 28:307–322.

Roda A, Mirasoli M, Michelini E, Di Fusco M, Zangheri M, Cevenini L, Roda B, Simoni P, (2016) Progress in chemical luminescence-based biosensors: A critical review. *Biosensors and Bioelectronics*, 76:164–179.

Rodríguez VM, Del Razo LM, Limón-Pacheco JH, Giordano M, Sánchez-Peña LC, Uribe-Querol E, Gutiérrez-Ospina G, Gonsebatt ME, (2005) Glutathione reductase inhibition and methylated arsenic distribution in Cd1 mice brain and liver. *Toxicological Sciences : An Official Journal of the Society of Toxicology*, 84:157–66.

Roquet N, Lu TK, (2014) Digital and analog gene circuits for biotechnology. *Biotechnology Journal*, 9:597–608.

Roquet N, Soleimany AP, Ferris AC, Aaronson S, Lu TK, (2016) Synthetic recombinase-based State machines in living cells. *Science*, 353:aad8559.

Rowland B, Purkayastha A, Monserrat C, Casart Y, Takiff H, McDonough KA, (1999) Fluorescence-based detection of lacZ reporter gene expression in intact and viable bacteria including Mycobacterium species. *FEMS Microbiology Letters*, 179:317–325.

Rubens JR, Selvaggio G, Lu TK, (2016) Synthetic mixed-signal computation in living cells. *Nature Communications*, 7:11658.

Saeidi N, Wong CK, Lo T-M, Nguyen HX, Ling H, Leong SSJ, Poh CL, Chang MW, (2011) Engineering microbes to sense and eradicate Pseudomonas aeruginosa, a human pathogen. *Molecular Systems Biology*, 7:521.

Saha RP, Samanta S, Patra S, Sarkar D, Saha A, Singh MK, (2017) Metal homeostasis in bacteria: the role of ArsR–SmtB family of transcriptional repressors in combating varying metal concentrations in the environment. *BioMetals*, 30:459–503.

Salis HM, Mirsky EA, Voigt CA, (2009) Automated design of synthetic ribosome binding sites to control protein expression. *Nature Biotechnology*, 27:946–950.

Salman M, Athar M, Waheed-uz-Zaman, Shafique U, Anwar J, Rehman R, Ameer S, Azeem M, (2012) Micro-determination of arsenic in aqueous samples by image scanning and computational quantification. *Analytical Methods*, 4:242.

Sanger CR, (1908) The Quantitative Determination of Arsenic by the Gutzeit Method. *Journal of the American Chemical Society*, 30:1041–1042.

Sansbury BM, Wagner AM, Nitzan E, Tarcic G, Kmiec EB, (2018) CRISPR-Directed In Vitro Gene Editing of Plasmid DNA Catalyzed by Cpf1 (Cas12a) Nuclease and a Mammalian Cell-Free Extract. *The CRISPR Journal*, 1:191–202.

Sattar A, Xie S, Hafeez MA, Wang X, Hussain HI, Iqbal Z, Pan Y, Iqbal M, Shabbir MA, Yuan Z, (2016) Metabolism and toxicity of arsenicals in mammals. *Environmental Toxicology and Pharmacology*, 48:214–224.

Schaerli Y, Gili M, Isalan M, (2014) A split intein T7 RNA polymerase for transcriptional AND-logic.

Nucleic Acids Research, 42:12322–12328.

Schiller CM, Fowler BA, Woods JS, (1977) Effects of arsenic on pyruvate dehydrogenase activation. *Environmental Health Perspectives*, Vol.19:205–207.

Schnütgen F, Doerflinger N, Calléja C, Wendling O, Chambon P, Ghyselinck NB, (2003) A directional strategy for monitoring Cre-mediated recombination at the cellular level in the mouse. *Nature Biotechnology*, 21:562–565.

Selifonova O, Burlage R, Barkay T, (1993) Bioluminescent sensors for detection of bioavailable Hg(II) in the environment. *Applied and Environmental Microbiology*, 59:3083–3090.

Selifonova O V., Eaton RW, (1996) Use of an ipb-lux fusion to study regulation of the isopropylbenzene catabolism operon of Pseudomonas putida RE204 and to detect hydrophobic pollutants in the environment. *Applied and Environmental Microbiology*, 62:778–783.

Shahid S, Hassan MI, Islam A, Ahmad F, (2017) Size-dependent studies of macromolecular crowding on the thermodynamic stability, structure and functional activity of proteins: in vitro and in silico approaches. *Biochimica et Biophysica Acta - General Subjects*, 1861:178–197.

Shalem O, Sanjana NE, Zhang F, (2015) High-throughput functional genomics using CRISPR-Cas9. *Nature Reviews Genetics*, 16:299–311.

Shaner NC, Campbell RE, Steinbach PA, Giepmans BNG, Palmer AE, Tsien RY, (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nature Biotechnology*, 22:1567–1572.

Sharma J, Mahajan R, Gupta V, (2010) Comparison and Suitability of Gel Matrix for Entrapping Higher Content of Enzymes for Commercial Applications. *Indian Journal of Pharmaceutical Sciences*, 72:223.

Sharma VK, Sohn M, (2009) Aquatic arsenic: Toxicity, speciation, transformations, and remediation. *Environment International*, 35:743–759.

Shemer B, Koshet O, Yagur-Kroll S, Belkin S, (2017) Microbial bioreporters of trace explosives. *Current Opinion in Biotechnology*, 45:113–119.

Sheth RU, Yim SS, Wu FL, Wang HH, (2017) Multiplex recording of cellular events over time on CRISPR biological tape. *Science*, 358:1457–1461.

Shetty RP, Endy D, Knight TF, (2008) Engineering BioBrick vectors from BioBrick parts. *Journal of Biological Engineering*, 2:5.

Shi W, Dong J, Scott RA, Ksenzenko MY, Rosen BP, (1996) The role of arsenic-thiol interactions in metalloregulation of the ars operon. *Journal of Biological Chemistry*, 271:9291–9297.

Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, Ueda T, (2001) Cell-free translation reconstituted with purified components. *Nature Biotechnology*, 19:751–755.

Shimomura O, Johnson FH, Saiga Y, (1962) Extraction, Purification and Properties of Aequorin, a Bioluminescent Protein from the Luminous Hydromedusan,Aequorea. *Journal of Cellular and Comparative Physiology*, 59:223–239.

Shin HJ, (2012) Agarose-gel-immobilized recombinant bacterial biosensors for simple and disposable on-site detection of phenolic compounds. *Applied Microbiology and Biotechnology*, 93:1895–1904.

Shin J, Noireaux V, (2010) Efficient cell-free expression with the endogenous E. Coli RNA polymerase and sigma factor 70. *Journal of Biological Engineering*, 4:8.

Shin J, Noireaux V, (2012) An E. coli cell-free expression toolbox: Application to synthetic gene circuits and artificial cells. *ACS Synthetic Biology*, 1:29–41.

Shis DL, Bennett MR, (2013) Library of synthetic transcriptional AND gates built with split T7 RNA polymerase mutants. *Proceedings of the National Academy of Sciences*, 110:5028–5033.

Siegal-Gaskins D, Tuza ZA, Kim J, Noireaux V, Murray RM, (2014) Gene circuit performance characterization and resource usage in a cell-free "breadboard." *ACS Synthetic Biology*, 3:416–425.

Siegfried K, Endes C, Bhuiyan AFMK, Kuppardt A, Mattusch J, van der Meer JR, Chatzinotas A, Harms H, (2012) Field testing of arsenic in groundwater samples of Bangladesh using a test kit based on lyophilized bioreporter bacteria. *Environmental Science and Technology*, 46:3281–3287.

Silver S, Phung LT, (2005) Genes and enzymes involved in bacterial oxidation and reduction of inorganic arsenic. *Applied and Environmental Microbiology*, 71:599–608.

Siuti P, Yazbek J, Lu TK, (2013) Synthetic circuits integrating logic and memory in living cells. *Nature Biotechnology*, 31:448–452.

Stark WM, (2017) Making serine integrases work for us. *Current Opinion in Microbiology*, 38:130–136.

Stocker J, Balluch D, Gsell M, Harms H, Feliciano J, Daunert S, Malik KA, van der Meer JR, (2003) Development of a set of simple bacterial biosensors for quantitative and rapid measurements of arsenite and arsenate in potable water. *Environmental Science and Technology*, 37:4743–4750.

Stoyanov J V., Browns NL, (2003) The Escherichia coli copper-responsive copA promoter is activated by gold. *Journal of Biological Chemistry*, 278:1407–1410.

Subach OM, Cranfill PJ, Davidson MW, Verkhusha V V., (2011) An enhanced monomeric blue fluorescent protein with the high chemical stability of the chromophore. *PLoS ONE*, 6:e28674.

Sun ZZ, Yeung E, Hayes C a., Noireaux V, Murray RM, (2014) Linear DNA for rapid prototyping of synthetic biological circuits in an escherichia coli based TX-TL cell-free system. *ACS Synthetic Biology*, 3:387–397.

Takahashi MK, Chappell J, Hayes CA, Sun ZZ, Kim J, Singhal V, Spring KJ, Al-Khabouri S, Fall CP, Noireaux V, Murray RM, Lucks JB, (2015a) Rapidly Characterizing the Fast Dynamics of RNA Genetic Circuitry with Cell-Free Transcription-Translation (TX-TL) Systems. *ACS Synthetic Biology*, 4:503–515.

Takahashi MK, Hayes CA, Chappell J, Sun ZZ, Murray RM, Noireaux V, Lucks JB, (2015b) Characterizing and prototyping genetic networks with cell-free transcription-translation reactions. *Methods*, 86:60–72.

Tan C, Saurabh S, Bruchez MP, Schwartz R, Leduc P, (2013) Molecular crowding shapes gene expression in synthetic cellular nanosystems. *Nature Nanotechnology*, 8:602–8.

Tang W, Liu DR, (2018) Rewritable multi-event analog recording in bacterial and mammalian cells. *Science*, 360:eaap8992.

Tong H-W, Mutlu BR, Wackett LP, Aksan A, (2014) Manufacturing of bioreactive nanofibers for bioremediation. *Biotechnology and Bioengineering*, 111:1483–93.

Trang PTK, Berg M, Viet PH, Van Mui N, Van Der Meer JR, (2005) Bacterial bioassay for rapid and accurate analysis of arsenic in highly variable groundwater samples. *Environmental Science and Technology*, 39:7625–7630.

UNITN iGEM, (2015) Interlab Measurement Study. http://2015.igem.org/Team:UNITN-Trento/Measurement. Accessed 15 Jun 2018

van der Meer JR, (2010) Bacterial Sensors: Synthetic Design and Application Principles. *Synthesis Lectures on Synthetic Biology*, 2:1–167.

van der Meer JR, Belkin S, (2010) Where microbiology meets microengineering: Design and applications of reporter bacteria. *Nature Reviews Microbiology*, 8:511–522.

Venturelli OS, Tei M, Bauer S, Chan LJG, Petzold CJ, Arkin AP, (2017) Programming mRNA decay to modulate synthetic circuit resource allocation. *Nature Communications*, 8:15128.

Virta M, Lampinen J, Karp M, (1995) A Luminescence-Based Mercury Biosensor. *Analytical Chemistry*, 67:667–669.

Volpetti F, Petrova E, Maerkl SJ, (2017) A Microfluidic Biodisplay. *ACS Synthetic Biology*, 6:1979–1987.

Wackwitz A, Harms H, Chatzinotas A, Breuer U, Vogne C, van der Meer JR, (2008) Internal arsenite bioassay calibration using multiple bioreporter cell lines. *Microbial Biotechnology*, 1:149–157.

Wang B, Barahona M, Buck M, (2013) A modular cell-based biosensor using engineered genetic logic circuits to detect and integrate multiple environmental signals. *Biosensors & Bioelectronics*, 40:368–76.

Wang B, Barahona M, Buck M, (2014) Engineering modular and tunable genetic amplifiers for scaling transcriptional signals in cascaded gene networks. *Nucleic Acids Research*, 42:9484–9492.

Wang B, Barahona M, Buck M, (2015) Amplification of small molecule-inducible gene expression via

tuning of intracellular receptor densities. *Nucleic Acids Research*, 43:1955–1964.

Wang B, Buck M, (2012) Customizing cell signaling using engineered genetic logic circuits. *Trends in Microbiology*, 20:376–384.

Wang B, Buck M, (2014) Rapid engineering of versatile molecular logic gates using heterologous genetic transcriptional modules. *Chem Commun*, 50:11642–11644.

Wang B, Kitney RI, Joly N, Buck M, (2011a) Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. *Nature Communications*, 2:508.

Wang Y, Yau YY, Perkins-Balding D, Thomson JG, (2011b) Recombinase technology: Applications and possibilities. *Plant Cell Reports*, 30:267–285.

Watstein DM, Styczynski MP, (2018) Development of a Pigment-Based Whole-Cell Zinc Biosensor for Human Serum. *ACS Synthetic Biology*, 7:267–275.

Way JC, Collins JJ, Keasling JD, Silver PA, (2014) Integrating biological redesign: Where synthetic biology came from and where it needs to go. *Cell*, 157:151–161.

Weber W, Fussenegger M, (2012) Emerging biomedical applications of synthetic biology. *Nature Reviews Genetics*, 13:21–35.

Webster DP, TerAvest M a, Doud DFR, Chakravorty A, Holmes EC, Radens CM, Sureka S, Gralnick J a, Angenent LT, (2014) An arsenic-specific biosensor with genetically engineered Shewanella oneidensis in a bioelectrochemical system. *Biosensors & Bioelectronics*, 62:320–4.

Wei Q, Nagi R, Sadeghi K, Feng S, Yan E, Ki SJ, Caire R, Tseng D, Ozcan A, (2014) Detection and spatial mapping of mercury contamination in water samples using a smart-phone. *ACS Nano*, 8:1121–1129.

Wen KY, Cameron L, Chappell J, Jensen K, Bell DJ, Kelwick R, Kopniczky M, Davies JC, Filloux A, Freemont PS, (2017) A Cell-Free Biosensor for Detecting Quorum Sensing Molecules in P. aeruginosa-Infected Respiratory Samples. *ACS Synthetic Biology*, 6:2293–2301.

WHO, (2005) Mercury in Drinking-water. *World Health Organization, Geneva*, 1–8.

WHO (ed), (2011) Guidelines for Drinking-water Quality, 4th edn. WHO Press, Geneva, Geneva

Wikipedia, (2018a) Sentinel species. https://en.wikipedia.org/wiki/Sentinel_species. Accessed 29 Dec 2017

Wikipedia, (2018b) John Scott Haldane. https://en.wikipedia.org/wiki/John_Scott_Haldane. Accessed 29 Dec 2017

Willsky GR, Malamy MH, (1980) Effect of arsenate on inorganic phosphate transport in Escherichia coli. *Journal of Bacteriology*, 144:366–374.

Wright O, Delmans M, Stan GB, Ellis T, (2015) GeneGuard: A modular plasmid system designed for biosafety. *ACS Synthetic Biology*, 4:307–316.

Wright O, Stan G-B, Ellis T, (2013) Building-in biosafety for synthetic biology. *Microbiology (Reading, England)*, 159:1221–35.

Wu CH, Le D, Mulchandani A, Chen W, (2009) Optimization of a whole-cell cadmium sensor with a toggle gene circuit. *Biotechnology Progress*, 25:898–903.

Wu F, Zhang Q, Wang X, (2018) Design of Adjacent Transcriptional Regions to Tune Gene Expression and Facilitate Circuit Construction. *Cell Systems*, 6:206–215.e6.

Wu M, Liang J, Tang J, Li G, Shan S, Guo Z, Deng L, (2017) Decontamination of multiple heavy metals-containing effluents through microbial biotechnology. *Journal of Hazardous Materials*, 337:189–197.

Wu YY, Culler S, Khandurina J, Van Dien S, Murray RM, (2015) Prototyping 1,4-butanediol (BDO) biosynthesis pathway in a cell-free transcription-translation (TX-TL) system. *BioRxiv*, 017814.

Xu C, Shi W, Rosen BP, (1996) The chromosomal arsR gene of Escherichia coli encodes a trans-acting metalloregulatory protein. *Journal of Biological Chemistry*, 271:2427–2432.

Yang L, Nielsen AAK, Fernandez-Rodriguez J, McClune CJ, Laub MT, Lu TK, Voigt CA, (2014) Permanent genetic memory with >1 byte capacity. *Nature Methods*, 11:1261–1266.

Yu B, Yang M, Shi L, Yao Y, Jiang Q, Li X, Tang LH, Zheng BJ, Yuen KY, Smith DK, Song E, Huang JD, (2012) Explicit hypoxia targeting with tumor suppression by creating an "obligate" anaerobic Salmonella Typhimurium strain. *Scientific Reports*, 2:436.

Zahir F, Rizwi SJ, Haq SK, Khan RH, (2005) Low dose mercury toxicity and human health. *Environmental Toxicology and Pharmacology*, 20:351–360.

Zhang F, Carothers JM, Keasling JD, (2012) Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nature Biotechnology*, 30:354–9.

Zhang F, Keasling J, (2011) Biosensors and their applications in microbial metabolic engineering. *Trends in Microbiology*, 19:323–9.

Zhang H, Lin M, Shi H, Ji W, Huang L, Zhang X, Shen S, Gao R, Wu S, Tian C, Yang Z, Zhang G, He S, Wang H, Saw T, Chen Y, Ouyang Q, (2014) Programming a Pavlovian-like conditioning circuit in Escherichia coli. *Nature Communications*, 5:3102.

Zhang H, Sheng Y, Wu Q, Liu A, Lu Y, Yin Z, Cao Y, Zeng W, Ouyang Q, (2013) Rational design of a biosensor circuit with semi-log dose-response function in Escherichia coli. *Quantitative Biology*, 1:209–220.

Zhang L, Wong MH, (2007) Environmental mercury contamination in China: Sources and impacts. Environ. Int. 33:108–121

Zheng L, Baumann U, Reymond JL, (2004) An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Research*, 32:e115.